UNIVERSIDAD AUTÓNOMA DE MADRID

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

Deciphering the structure of CAD: structural and functional characterization of the human aspartate transcarbamoylase domain

Alba Ruiz Ramos

Madrid, 2016

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

FACULTAD DE CIENCIAS

UNIVERSIDAD AUTÓNOMA DE MADRID

Deciphering the structure of CAD: structural and functional characterization of the human aspartate transcarbamoylase domain

Alba Ruiz Ramos

Thesis Directors

Dr. Santiago Ramón Maiques
Dr. María Moreno Morcillo
Structural Bases of Genome Integrity Group
Structural Biology and Biocomputing Programme
Spanish National Cancer Research Centre (CNIO)

Doctoral thesis submitted to the Universidad Autónoma de Madrid for the degree of PhD by M. Sci. in Biology,

Alba Ruiz Ramos

Centro Nacional de Investigaciones Oncológicas

Dr. Santiago Ramón Maiques, head of the Structural Bases of Genome Integrity Group in the Spanish National Cancer Research Centre (CNIO), and Dr. María Moreno Morcillo, Postdoctoral Researcher in the same group,

CERTIFY

That Ms Alba Ruiz Ramos, Master in Biology by the University of Córdoba, has completed her Doctoral Thesis **"Deciphering the structure of CAD: structural and functional characterization of the human aspartate transcarbamoylase domain"** and meets the necessary requirements to obtain the PhD Degree in Molecular Biosciences. To this purpose, she will defend her Doctoral Thesis at the Universidad Autónoma de Madrid. The Thesis has been carried out under our direction and hereby we authorize it to be defended to the appropriate Thesis Tribunal.

We hereby issue this certificate in Madrid, 19th of July 2016.

Santiago Ramón Maiques

PhD Thesis Director

María Moreno Morcillo

PhD Thesis Co-director

ACKNOWLEDGMENTS

Esta tesis la defiendo yo. Pero siento que el trabajo que aquí se presenta no es solo mío, hay mucha gente que me ha ayudado a que este barco llegue a buen puerto.

El primero al que tengo que darle las gracias es a mi director, Santiago. Gracias por confiar en mi cuando llegué de Córdoba y pensaste que lo de cristalizar proteínas se me podría dar bien. Casi aciertas. Gracias por todo, por enseñarme muchísimo de proteínas, purificaciones, cristalización, etc. Por ser mi becario, y mano derecha cuando la mía no funcionaba. Gracias por tu paciencia, porque en estos cinco años la he liado bien parda haciendo los experimentos y siempre estabas ahí para sacarme del lio. Pero, sobre todo, por hacerme un hueco en el despacho y nunca quejarte por tenerme de compañera, haciendo ruido y hablando conmigo misma sin parar. Sé que ha sido duro, pero nos hemos divertido mucho. Yo, por lo menos.

Después de Santiago, están mis compañeros de laboratorio. Los que fueron y los que son. Cómo no quiero llorar ni ponerme sentimental, que yo no soy así, os doy las gracias a todos a la vez. Gracias por ayudarme en todo, por enseñarme todo lo que sé en el laboratorio pero, sobre todo, por convertiros en mis amigos. Para mí, eso es mucho más importante que la ciencia. Ara, Nada, Marija, Leyre, Paco y María (y demás) esto es vuestro también!!! Nada de esto estaría aquí escrito sino es por vosotros. Gracias, gracias y gracias.

A mi Familia. Gracias por aguantar mis ausencias sin un reproche, por escuchar al teléfono todas mis quejas y lloros, por alegrarse mucho cuando les decía que algo (que ellos no entienden ni de lo que se trata) había funcionado. Yo soy yo por y para ellos. Gracias por el apoyo constante, porque como algún amigo me dijo antes, “ser doctorando es jodido, pero ser familiar de doctorando, lo es más”. Esto va por vosotros.

A todos lo que sientan que esto es un poco suyo, también muchas gracias. A los amigos de todos lados, a los que han sufrido que esté lejos y con las proteínas, a los que organizaron eventos y yo no estaba. Gracias por entender dónde tenía que estar y esperarme con una sonrisa.

Y por último, a Carlos. Mi persona favorita del mundo mundial. Cada letra de este documento es tan tuya como mía. Gracias por ayudarme en la distancia, por las palabras de apoyo, luchar por nosotros y por creer en mi siempre, aunque a veces era complicado dadas las circunstancias. Te quiero mucho. Se acabó el sufrir, ahora empieza lo bueno!!

ABSTRACT

Aspartate transcarbamoylase (ATCase) catalyzes the synthesis of N-carbamoyl-L-aspartate from carbamoyl phosphate and aspartate in the second step of the *de novo* biosynthesis of pyrimidines. In prokaryotes and plants, the first three activities of the pathway, namely carbamoyl phosphate synthetase (CPSase), ATCase and dihydroorotase (DHOase), are encoded as distinct proteins that function independently or in non-covalent association. In turn, in animals, CPSase, ATCase and DHOase are part of a 243 kDa multifunctional polypeptide named CAD whose up-regulation is essential for normal and tumor cell proliferation. Although the structures of numerous prokaryotic ATCases have been determined, there is no structural information about any eukaryotic ATCase. In fact, the only detailed structural information about CAD is that it self-assembles through its ATCase domain forming hexameric particles of 1.5 MDa. In this study, we report the cloning, expression and purification of the ATCase domain of human CAD. The recombinant protein, expressed in bacteria and purified with good yields, forms homotrimers in solution. Crystallization experiments, both in the absence and presence of the inhibitor PALA, yielded small crystals that diffracted X-rays to 2.1 Å resolution using synchrotron radiation. We have determined the structure of the human ATCase domain, the first example of a eukaryotic ATCase, confirming the overall similarity with bacterial homologues. Surprisingly, the enzymatic and biochemical characterization of human ATCase showed that this domain exhibits cooperativity effects that reduce the affinity for the anti-tumoral drug PALA. These results were unexpected since the prokaryotic ATCase catalytic trimers lack cooperativity between active sites. Combining structural, mutagenic and biochemical analysis we identified key structural elements in human ATCase for the necessary regulation and transmission of conformational changes leading to cooperativity between subunits. Mutation of one of these elements, the active site residue R2024, was recently found to cause the first non-lethal CAD-deficit. We reproduced this mutation in the recombinant human ATCase domain and measured its effect. Our data suggest that this arginine is part of a molecular switch that regulates the conformational equilibrium between low and high affinity states of the subunits for the ligands. To further understand the role of ATCase trimers in the assembly of CAD particles, we produced a bifunctional construct harbouring the DHOase and ATCase domains of human CAD together with the 91 amino acid interlinking sequence. We demonstrated that this construct assembles into hexamers in solution and that site-specific mutations preventing either the dimerization of DHOase or the trimerization of ATCase, results in DHOase-ATCase trimers and dimers respectively. These results indicate that the ATCase and DHOase domains form the structural core of CAD particles.

RESUMEN

La aspartato transcarbamilasa (ATCasa) cataliza la síntesis de N-carbamil-L-aspartato a partir de carbamil fosfato y aspartato en el segundo paso de la síntesis *de novo* de pirimidinas. En procariotas y en plantas, las primeras tres actividades de esta ruta, llamadas carbamil fosfato sintetasa (CPSasa), ATCasa y dihidroorotasa (DHOasa), están codificadas como proteínas que funcionan de forma independiente o asociadas de forma no covalente. En cambio en animales CPSasa, ATCasa y DHOasa son parte de un polipéptido multifuncional de 243 kDa llamado CAD, cuya activación es esencial para la proliferación de células normales y tumorales. Aunque se conoce la estructura de numerosas ATCasas procariotas, la única información estructural detallada sobre CAD es que se auto-ensambla a través de su dominio ATCasa formando partículas hexaméricas de 1.5 MDa. En este estudio describimos la clonación, expresión y purificación del dominio ATCasa de CAD humana. La proteína recombinante, expresada en bacterias y purificada con buen rendimiento, forma homo-trímeros en solución. Los experimentos de cristalización, tanto en ausencia como en presencia del inhibidor PALA, produjeron cristales que difractan los rayos X a una resolución de 2.1 Å usando radiación de sincrotrón. Hemos determinado la estructura del dominio ATCasa humano, siendo la primera estructura de una ATCasa eucariota, confirmando la similitud general con los homólogos bacterianos. Sorprendentemente, la caracterización enzimática y bioquímica de la ATCasa humana mostró que este dominio exhibe propiedades cooperativas que reducen la afinidad por el compuesto anti-tumoral PALA. Combinando el análisis estructural, mutagénico y bioquímico hemos identificado elementos estructurales claves para la transmisión de cambios conformacionales que permiten la cooperatividad entre subunidades. La mutación de uno de estos elementos, el residuo del centro activo R2024, se ha descrito recientemente como la causa del primer déficit no letal de CAD. Hemos reproducido esta mutación en la ATCasa humana recombinante y medido su efecto. Nuestros datos sugieren que esta arginina es parte de un interruptor molecular que regula el equilibrio conformacional entre estados de baja y alta afinidad de las subunidades por los ligandos. Para comprender mejor el papel de los trímeros de ATCasa en la arquitectura de CAD, produjimos una construcción bifuncional que contiene los dominios DHOasa y ATCasa humanos. Hemos demostrado que esta construcción forma hexámeros en solución y que mutaciones específicas que previenen la dimerización de los dominios DHOasa o la trimerización de la ATCasa resultan respectivamente en la formación de trímeros o dímeros de DHOasa-ATCasa. Estos resultados prueban que los dominios ATCasa y DHOasa forman el esqueleto central de las partículas de CAD.

TABLE OF CONTENTS

ABBREVIATIONS LIST

Asp	Aspartate
ATCase	Aspartate transcarbamoylase
AUC	Analytical ultracentrifugation
BSA	Bovine serum albumine
CAD	Carbamoyl-phosphate synthetase II, Aspartate transcarbamylase, and Dihydroorotase
CAP	Catabolite Activator Protein
CAsp	Carbamoyl aspartate
CHO	Chinese hamster ovary
CP	Carbamoyl-phosphate
CPSase	Carbamoyl-phosphate synthetase
DHOase	Dihydroorotase
DHOdase	Dihydroorase dehydrogenase
DMEM	Dulbecco´s modified Eagle´s medium
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
FF	Fast flow
FPLC	Fast protein liquid chromatography
GLNase	Glutaminase
GST	Glutathione S-transferase
HEK293	Human embryonic kidney 293
IC_{50}	Concentration giving half-maximal inhibition
IPTG	Isopropyl ß-D-1-thiogalactopyranose
ITC	Isothermal titration calorimetry
K_D	Dissociation constant
K_M	Michaelis-Menten constant
LB	Luria Bertani growth media
MALS	Multi-angle light scattering
MAPK	Mitogen-activated protein kinase
MBP	Maltose binding protein
MS	Mass spectrometry
MW	Molecular weight
ODCase	Orotidylate decarboxylase
OD	Optical density
OMP	Orotidine 5-phosphate; orotidylate
OPRT	Orotate phosphoribosyl transferase
PAGE	Polyacrylamide gel electrophoresis
PALA	N-phosphonacetyl-L-aspartate

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data base
PEI	Polyethyleneimine
PMSF	Phenylmethylsulfonyl fluoride
PRPP	5-phosphoribosyl-1-pyrophosphate
RPM	Revolution per minute
SEC	Size-exclusion chromatography
SDS	Sodium dodecyl sulphate
Suc	Succinate
T_M	Melting temperature
UMPS	UMP synthase
V^{app}	Apparent limiting rate
WT	Wild-type

INTRODUCTION

1. *De novo* synthesis of pyrimidine nucleotides

Purine and pyrimidine nucleotides are the activated precursors of nucleic acids. As such, they constitute the building blocks for the replication and storage of information in the DNA and for the transcription of the genetic information into messenger RNA. A large fraction of the nucleotide pool is also present in the form of ribosomes (ribosomal RNA), which accumulate more than 80 % of the total nucleic acids of rapidly growing mammalian cells (Bremer and Dennis, 1996, Lodish et al., 2000, Lane and Fan, 2015). Purine nucleotides carry out a number of other important functions, such as acting as the cellular currency of energy and as donors of phosphoryl groups by protein kinases (ATP and GTP), as enzymatic cofactors, or as second messengers (cyclic AMP and GMP) in signal-transduction pathways (Cooper, 2000). On the other hand, pyrimidines nucleotides play a central role in the activation of sugars (UDP-sugars) for the formation of glycogen and for protein glycosylation (Bülter and Elling, 1999). Pyrimidines also form an important nucleotide-lipid conjugate, CDP-diacilglycerol, a branch point precursors of phospholipids with great importance for the assembly of the plasmatic membrane in eukaryotes. Finally, uridine nucleotides are also described to act as extracellular messengers regulating a variety of physiological processes (Connolly and Duley, 1999) (Figure 1).

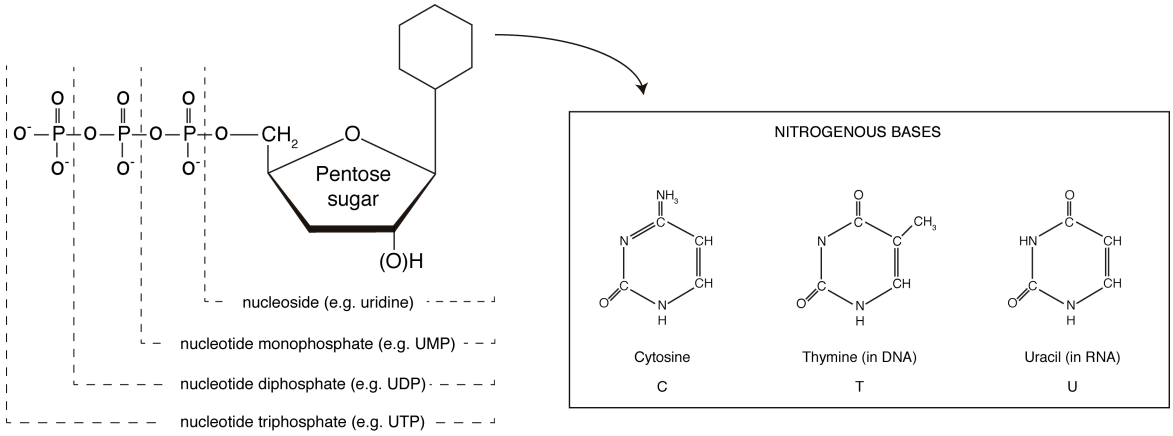

Figure 1. Pyrimidine nucleotides. Schematic representation of the most common variants of pyrimidine nucleotides.

Cells can obtain nucleotides through two different metabolic pathways. On the one hand, purine and pyrimidine bases, either released during the turnover of nucleic acids and nucleotides or obtained from the diet, can be salvaged and recycled by the cell. In these salvage pathways, the nucleotide bases (e.g. uracil) are reconnected to a ribose sugar activated in the form of 5-phosphoribosyl-1-pyrophosphate (PRPP) by specific phosphoribosyl transferases (PRTs). In addition, nucleosides (e.g. uridine) can be phosphorylated by specific kinases into the phosphorylated form. On the other hand, nucleotides can be synthesized *de novo* (from scratch) using simple compounds. Whereas

purine bases are made piece by piece directly onto a ribose structure, in *de novo* synthesis of pyrimidines the nitrogenous base is pre-assembled from bicarbonate, aspartate and glutamine, and later, attached to the activated 5-ribose unit, PRPP (Berg et al., 2002).

In this study, we focused our attention in *de novo* pathway for pyrimidine synthesis (Figure 2). This enzymatic route starts with the production of carbamoyl phosphate (CP) from bicarbonate, ammonia and two molecules of ATP, a reaction catalyzed by the carbamoyl phosphate synthetase (CPSase). In animals, there are two types of CPSases physically separated in different cell compartments. One CPSase, named type I, localizes in the mitochondria of hepatocytes and detoxifies ammonia by making CP in the urea cycle (Figure 2) (Nyunoya et al., 1985, Tramell and Campbell, 1970). The second CPSase (type II) is cytosolic, and catalyzes the dedicated production of CP for the synthesis of pyrimidines (Simmer et al., 1990). This enzyme synthesizes CP in four partial reactions (Anderson and Meister, 1966). First, bicarbonate is phosphorylated to carboxyphosphate, which reacts with ammonia producing carbamate and inorganic phosphate. Then, carbamate is phosphorylated to yield CP. CPSase primarily uses glutamine as the source of ammonia. The hydrolysis of the side chain of glutamine to produce ammonia is carried out by a glutaminase (GLNase) enzymatic activity associated with the CPSase (Simmer et al., 1990). The next step in *de novo* synthesis of pyrimidines is catalyzed by aspartate transcarbamoylase (ATCase), which condensates CP and aspartate (Asp) to form carbamoyl aspartate (CAsp) and Pi. Then, CAsp cyclizes to form dihydroorotate by the action of the zinc-metalloenzyme dihydroorotase (DHOase).

The next step in the pathway is the oxidation of the dihydroorotate to orotate by the enzyme dihydroorotate dehydrogenase (DHOdase). In eukaryotes, this enzyme is located at the inner mitochondrial membrane and couples the synthesis of pyrimidines to the mitochondrial electron transport chain (Bremer and Dennis, 1996, Rawls et al., 2000). Then, orotate is attached to the activated form of the ribose, PRPP, by the enzyme orotate phosphoribosyl transferase (OPRTase), to form the first pyrimidine nucleotide, orotidine 5-phosphate (or orotydylate; OMP). OMP is then decarboxylated to form uridine 5-phosphate (uridylate; UMP) by the orotidylate decarboxylase (ODCase), considered one of the most catalytically efficient enzymes known (Radzicka and Wolfenden, 1995).

UMP is a major pyrimidine nucleotide, but it needs to be phosphorylated by ATP-dependent kinases for the formation of the other major pyrimidines nucleotides. Thus, UMP is phosphorylated first to UDP by the UMP/CMP kinase, and later converted to UTP by the action of the nucleotide diphosphate kinase (NDK), an enzyme with a broad specificity that acts both on purines and pyrimidines. UTP is a precursor of RNA and can be considered the final product of *de novo* synthesis. The other pyrimidine for RNA synthesis, CTP, is obtained by the CTP synthetase, which uses ATP to activate a carbonyl oxygen of UTP, and then, replaces the phosphate by ammonia obtained by the hydrolysis of glutamine.

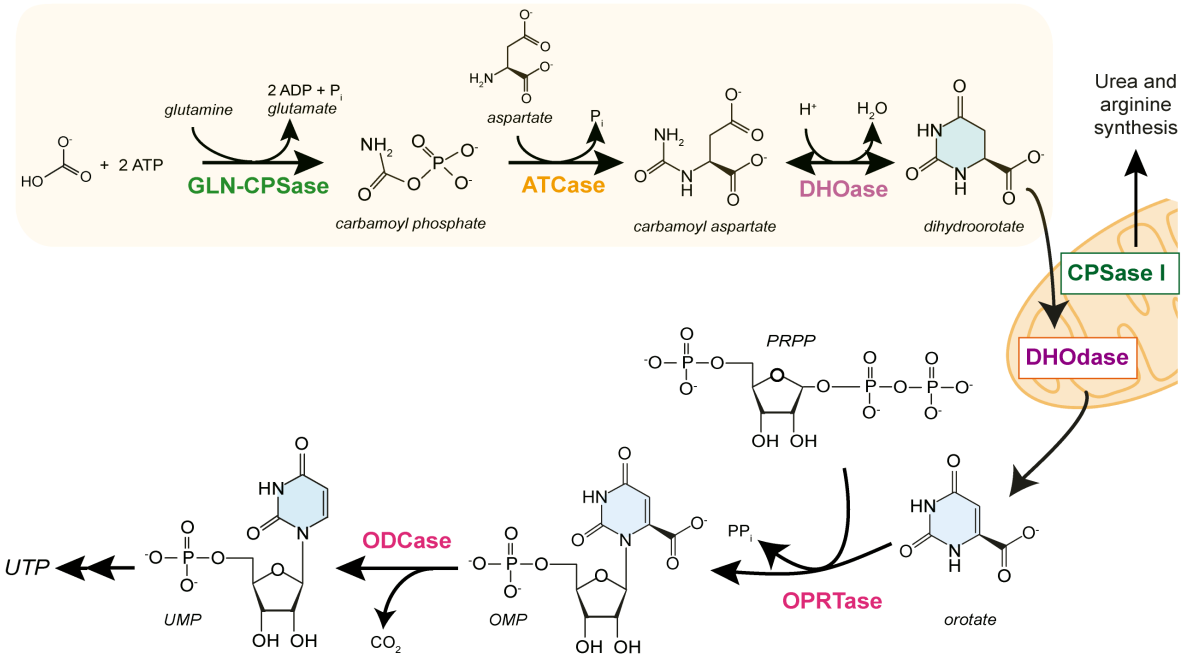

Figure 2. De *novo* biosynthesis of pyrimidines pathway in animals. *De novo* synthesis of UMP involves six reactions. In animals, the first three steps, highlighted in a yellow box, are catalyzed by the multifunctional protein CAD. The fourth step is catalyzed by a dehydrogenase located in the inner mitochondrial membrane. The CPSase activity is the rate-limiting step of the pathway and is also allosterically inhibited by the final product, UTP, and activated by PRPP. The CPSase I is also shown in the mitochondrion, despite it is not involved in pyrimidine biosynthesis in animals.

Finally, the synthesis of DNA requires the reduction of the 2'-hydroxyl group of the nucleotides. These precursors are formed by the ribonucleotide reductase, an enzyme responsible for the reduction of the four di-phospho nucleotides to deoxyribonucleotides. In addition, uracile nucleotides need to be dimethylated to thymidine prior to its incorporation to DNA. This final step is carried out by the thymidylate synthetase that converts dUMP to dTMP, which is later phosphorylated by thymidylate kinase and NDK to form dTTP.

2. Multifunctional proteins in *de novo* synthesis of pyrimidines

The six steps in *de novo* pathway leading to the synthesis of UMP are fully conserved in all living organisms. However, there is extensive diversity in the arrangement of the genes encoding the enzymes that catalyze these reactions, in the organization of the enzymes and in the regulation of the pathway (Jones, 1980). In prokaryotes, plants and most protozoa, the six enzymatic activities leading to the formation of UMP are encoded as distinct proteins, that either work alone or forming non-covalent complexes. In contrast, in animals, the six enzymes are encoded within just three polypeptides (Figure 3).

In 1971, the studies in mouse spleen carried out by Nicholas Hoogenraad and colleagues showed that the first two enzymes in *de novo* pathway, GLNase-dependent CPSase and ATCase, co-purified as a complex (Hoogenraad et al., 1971). Similar results had been previously described in yeast (Lue and Kaplan, 1969) and in molds (Williams et al., 1970). Mary Ellen Jones also described that the complex in rat ascites cells additionally included the DHOase activity (Shoaf and Jones, 1971). The nature of this complex was not

truly understood until it was purified to homogeneity by George Stark and colleagues from a hamster cell strain treated with the ATCase inhibitor PALA (N-phosphonacetyl-L-aspartate) (Coleman et al., 1977) (see details in BOX 1). When the complex was isolated, the CPSase, ATCase and DHOase activities were found to be fused into a single polypeptide chain with a molecular mass of over 240 kDa (Figure 3A). The isolated protein was a mixture of different oligomeric forms, mainly trimers and hexamers, composed of multiple copies of the protein. The association of these three enzymes into a multifunctional polypeptide was subsequently documented in every animal species investigated and received various names, such as complex A or pyr1-3, *rudimentary* (in drosophila), or *perplexed* (zebrafish). Nowadays, this chimeric multi-functional protein is commonly known as CAD, an acronym derived from the initials of the three activities encoded within the polypeptide: (GLNase)CPSase, ATCase and DHOase.

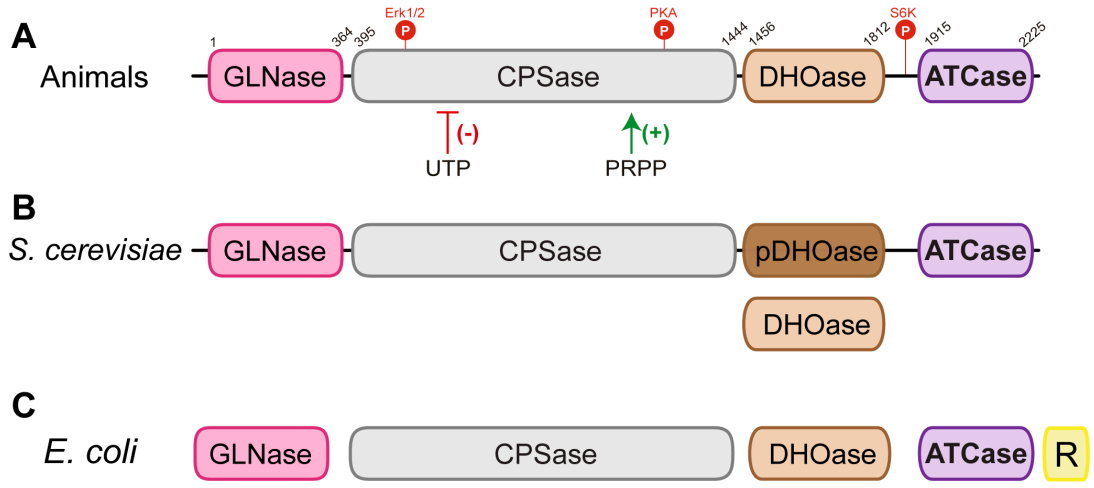

Figure 3. Organization of the three first enzymes of *de novo* pathway. (A) In animals, the first three steps are catalysed by the multifunctional protein CAD, shown schematically with each functional domain in a different colour. The CPSase activity is allosterically inhibited by UTP (in red) and activated by PRPP (in green). The phosphorylation sites and the residue numbers of the different domains are shown. **(B)** In fungi, e.g. *S. cerevisiae*, CPSase and ATCase are part of a multifunctional CAD-like protein that contains an inactive DHOase domain (pDHOase). The DHOase activity is provided by an independently encoded protein. **(C)** In bacteria, archaea and plants the enzymes function independently, although they can form either covalently or non-covalently linked complexes. In the case of *E. coli*, the ATCase forms a non-covalent complex with regulatory subunits (shown in yellow) that can be allosterically regulated.

Further studies led among others by the groups of David Evans, Jeffrey Davidson and Elizabeth Carrey helped to assign the enzymatic activities to independent functional domains within the CAD polypeptide. These studies relied on the extraordinary susceptibility of the protein to cleavage by specific proteases (Carrey and Hardie, 1988, Hemmens and Carrey, 1994, Davidson et al., 1981, Kim et al., 1992, Mally et al., 1981). The selective proteolytic cleavage and subsequent isolation of fragments retaining the enzymatic activities provided strong evidence for the domain structure of CAD, later verified by protein and gene sequencing (Carrey and Hardie, 1988) (Figure 3A).

BOX 1. PALA and the discovery of CAD. The discovery of CAD tracks back to the 70's, and is intimately ligated to the synthesis of PALA. Following Linus Pauling's reasoning that analogs of the transition state would be potent reversible inhibitors of any enzymatic activity, Kim Collins and George Stark synthesized PALA, postulating that it would be a transition state analogue of ATCase (Collins and Stark, 1971). Indeed, PALA turn out to be a potent and specific inhibitor of *E. coli* ATCase.

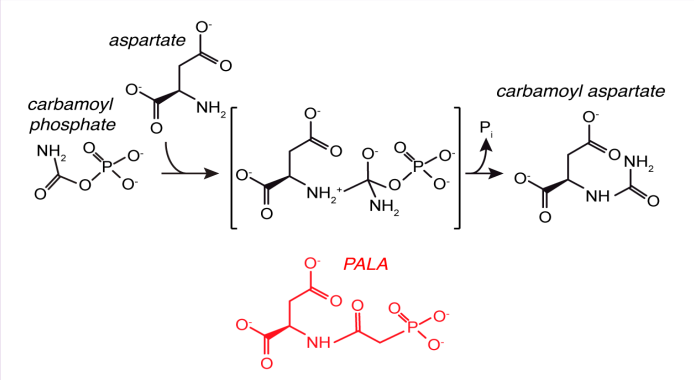

Figure 4. ATCase catalytic reaction and PALA. The tetrahedral intermediate state in the reaction is shown in brackets (Gouaux et al., 1987). PALA, shown in red, combines features of both substrates and resembles to some extent the transition state.

They also observed that PALA penetrated easily in mammalian cells cultures and mice, blocking de novo biosynthesis of pyrimidines and arresting cell proliferation. This assay demonstrated for the first time that *de novo* synthesis of pyrimidines was essential for cell growth and proliferation. However, a fraction of PALA-treated hamster CHO cells overcome the inhibitory effect by producing up to 150-fold higher levels of ATCase (Kempe et al., 1976). Patrick Coleman and Parker Suttle used these PALA-resistant cells to isolate the mammalian ATCase, realizing that it co-purified forming a very stable complex with the CPSase and DHOase activities. Later on, they would discover that the three enzymes were covalently linked in a single giant polypeptide that they named CAD (Coleman et al., 1977) and that self-associated forming different oligomers, mainly trimers and hexamers. Further research by the group of Stark demonstrated that the PALA-resistant CHO cells underwent a massive amplification (up to 100-fold) of the cad gene (Wahl et al., 1979). This was the first evidence of gene amplification in mammalian cells, a common hallmark of cancer cells.

Interestingly, in fungi, GLN-CPSase and ATCase are also fused into a ~240 kDa CAD-like polypeptide (e.g. named *URA2* in *Saccharomyces cerevisiae*) (Denis-Duphil and Kaplan, 1976), which contains a defective DHOase-like domain (Williams et al., 1970, Souciet et al., 1989). This activity is provided by a separated gene (*URA4* in *S. cerevisiae*) encoding an independent and catalytically active DHOase (Guyonvarch et al., 1988).

CAD is not the only multi-functional protein found in *de novo* pyrimidine pathway in animals. In a similar manner, the reactions 5 and 6 in the pathway, catalyzed by OPRTase

and ODCase, depend on a bifunctional protein, the UMP synthase, which has evolved from the fusion of the two independent genes. This fusion also occurs in certain pathogens, although interestingly, in some instances, the order of the functional domains within the polypeptide is the reverse to that observed in animals (Traut and Jones, 1996). Enzymes catalyzing sequential reactions are commonly organized into macromolecular complexes. The assembly of enzymes into complexes offers advantages for improving the metabolic efficiency (see details in BOX 2).

BOX 2. What are the advantages of a multi-functional protein? Multienzyme complexes can be composed of individual proteins that interact non-covalently, or they might result of a single polypeptide with several functional domains, such as CAD. The co-localization of different enzymes, by forming a multienzyme, favors the transfer of intermediates between the active sites without diffusion to the solvent, allowing the global coordination of the different activities. It also favors the forward direction of the reversible reactions, since products are taken by the next activity in the pathway, and also avoids the interference with other metabolic pathways in the cell. In eukaryotes, there are examples of these multifunctional polypeptides that must have been evolved by gene recombination. The number of the functional domains found in these multifunctional polypeptides can vary; for example, there are two in UMPS (Jones, 1980), two in tryptophan synthetase (Dunn et al., 2008), two in acetyl-CoA carboxylase (Tran et al., 2015, Barber et al., 2005), three in CAD (four if we consider separately the GLNase domain), four in *E. coli* fatty acid oxidation α-subunit (Yang and Schulz, 1983), five in the AROM complex in fungi (Hawkins et al., 1993), seven in animal fatty acid synthetase (Smith, 1994, Witkowski et al., 1991, Wakil, 1989), and up to ten in polyketide synthase (Dutta et al., 2014). The conservation of some of these fusions from yeast to humans implies that the multi-functional proteins may have additional selective advantages over the monofunctional prokaryotic counterparts and might reflect additional layers of complexity. What are these advantages? Encoding several proteins within a single transcript is the equivalent to a bacterial operon, a polycistronic messenger, with multiple translation initiation sites (Davidson et al., 1993, Coleman et al., 1977). Thus, gene fusions facilitate the production of proteins simultaneously and in stoichiometric amounts. The covalent linkage must enhance the association between the functional domains, which would otherwise diffuse freely and would depend on strong and highly specific interactions to assemble as a complex. These favored interactions might also increase the integrity of individual components, that would be unstable out of the complex. Furthermore, it is suggested that the order in which the genes are fused could be relevant for the folding and ordered assembly of the multifunctional complex (Marsh et al., 2013).

3. Cell growth, proliferation and regulation of *de novo* pathway by CAD

Pyrimidines are required for a wide variety of biological processes and are constantly made *de novo* in all cells. However, in *de novo* pathway each molecule of nucleotide is produced at high-energy price compare to salvage pathways, which are much more energetically efficient. This explains that the rate of *de novo* synthesis is reduced when exogenous uridine to feed the salvage pathway is supplied (Hoogenraad and Lee, 1974). In general, fully differentiated and quiescent cells obtain the supply of pyrimidines largely through salvage pathways, whereas the activity of *de novo* synthesis is in a low basal level. However, in rapidly growing and dividing cells the recycling route is not sufficient and the flux through *de novo* pathway is up-regulated to fuel the high demand of nucleotides needed for RNA and DNA synthesis, ribosome ribogenesis and other anabolic processes (Evans and Guy, 2004).

The mechanism of enzymatic regulation of *de novo* pyrimidine synthesis is different in prokaryotes and in higher eukaryotes (Figure 2 and 3). For instance, in bacteria, CP is the product of the first reaction in the pathway to make UMP, but it is also the substrate of others transcarbamoylases, such as ornithine transcarbamoylase (OTCase), which is involved in the biosynthesis of arginine. Thus, the first committed step to the biosynthesis of pyrimidines in prokaryotes is that catalyzed by ATCase, which is feedback inhibited by the final products of the pathway, CTP and UTP, and activated by the allosteric effector ATP (Kantrowitz, 2012) (Figure 5). Contrary, in animals, the CPSase domain of CAD (CPSase II) catalyzes the dedicated production of CP for the synthesis of pyrimidines, being the limiting rate reaction of pathway. Thus, CPSase II is the control point of the pathway, being feedback inhibited by UTP and allosterically activated by PRPP (Liu et al., 1994) (Figure 3). Hence, unlike some bacteria, in animals, the ATCase domain of CAD is not subjected to a direct regulation by allosteric effectors.

The allosteric regulation of CPSase II by UTP and PRPP is further modulated by phosphorylation through two signaling cascades (Figure 3A). The mitogen-activated protein (MAP) kinase cascade (Erk1/2) phosphorylates CAD in residue Thr456, making CAD insensitive to feedback inhibition by UTP, converting UTP into a mild activator, and increasing the susceptibility of CAD to the activation by PRPP (Graves et al., 2000). CPSase II is also phosphorylated by cyclic AMP-dependent protein kinase (PKA) at Ser1406 (Carrey et al., 1985), resulting in a loss of feedback inhibition by UTP and in a decreased sensitivity to PRPP (Sahay et al., 1998). The phosphorylations by MAPK/ERK and PKA appear to be sequential and mutually exclusive during the life cycle of the cell, modulating CAD activity to satisfy the needs of pyrimidines during division (Sigoillot et al., 2003, Sigoillot et al., 2002). In this manner, when the resting cells are stimulated to proliferate, the MAP kinase cascade up-regulates CAD activity by phosphorylating Thr456 just prior to the S-phase of the cell cycle, which results in an increased in the flux of the pathway and in the pyrimidine synthesis.

When cells come out of S-phase, Thr456 is dephosphorylated, and PKA phosphorylates Ser1406, turning CAD insensitive to the activation by PRPP and returning the flux of the pathway to resting levels.

Two recent independent studies have also demonstrated that CAD activity is also stimulated by phosphorylation through the mTORC1 (mechanistic target of rapamycin complex 1) pathway (Robitaille et al., 2013, Ben-Sahra et al., 2013) (Figure 3A). In response to nutrients, protein S6 kinase 1 (S6K) phosphorylates residue Ser1859 at the linker between the DHOase and ATCase domains (Robitaille et al., 2013, Ben-Sahra et al., 2013). This phosphorylation site is conserved among CAD sequences in vertebrates (Ben-Sahra et al., 2013), and can also be phosphorylated *in vitro* by PKA (Carrey et al., 1985). The mTORC1-dependent phosphorylation up-regulates the activity of CAD promoting pyrimidine synthesis and S phase progression. This phosphorylation also induces clustering of CAD in the cytoplasm and varies the sedimentation profile of the protein, suggesting that it promotes changes in the oligomerization of the protein (Robitaille et al., 2013).

CAD is also reported to autophosphorylate at residue Thr1037 (Carrey et al., 1985, Sigoillot et al., 2002). This leads to an increased sensitivity to UTP and loss of activation by PRPP. However, the level of autophosphorylation remains constant throughout the cell cycle and therefore, its significance is unclear (Sigoillot et al., 2003).

The activity of CAD is also regulated by its own synthesis and degradation. At the transcriptional level, CAD is regulated by Myc. The transcription of *cad* gene increases at the G_1/S-phase boundary, when quiescent cells enter the proliferative cell cycle (Boyd and Farnham, 1997). This increase correlates with the c-myc expression peaks in dividing cells, which directly binds to the *cad* gene promoter (Huang and Graves, 2003). Contrary, c-myc expression is suppressed as cells exit the cell cycle or terminally differentiate, reducing the transcription of *cad* gene (Bush et al., 1998). On the other hand, CAD is rapidly broken down by caspase-mediated degradation during cell death, since it carries two caspase-3 cleavage sites located in CPSase domain (Huang and Graves, 2003).

4. CAD in cancer and in other diseases

The up-regulation of CAD activity is essential for cell proliferation, particularly of tumors and neoplastic cells (Aoki and Weber, 1981, Sigoillot et al., 2004). Early studies demonstrated that the ATCase activity is 2 to 4-fold higher in hepatoma nodules than in adjacent liver tissue (Calva et al., 1959). Higher increments (up to 10-fold) in CPSase II activity were described in hepatomas when compared with normal liver (Aoki and Weber, 1981). The activity of *de novo* pathway is increased 8-fold when BHK (baby hamster kidney) cells enter in exponential growth and drops sharply when the cells are confluent (Sigoillot et al., 2002). It has been also shown that the intracellular concentration of CAD is increased 3.5 to 4-fold in MCF7 human breast cancer cells compared with normal cells, and remains

persistently activated along the cell-cycle (Sigoillot et al., 2004). This persistent activation correlates with the imbalance in the relative levels of Erk1/2 kinase and PKA activities in tumor cells (Sigoillot et al., 2004).

The up-regulation of CAD activity in malignant cells turns this multienzymatic protein into a potential anti-tumoral target for pharmacological intervention (Christopherson and Lyons, 1990, Christopherson et al., 2002, Löffler et al., 2005). However, only one anti-CAD compound, PALA, has been used in the clinic for cancer treatment (Collins and Stark, 1971, Johnson et al., 1976, Grem et al., 1988, Swyryd et al., 1974). A year before the discovery of CAD, in 1976, George Stark and colleagues tested the use of PALA as an antitumor agent in mice, with initial spectacular results (Johnson et al., 1976). They proved that PALA reduces the synthesis of pyrimidines (Yoshida et al., 1974, Swyryd et al., 1974) and effectively blocks the growth of different types of murine solid tumors (Grem et al., 1988, Tsuboi et al., 1977). In 1978, PALA was introduced into clinical trials (Grem et al., 1988) but failed to prove its efficacy as a single chemotherapeutic agent (Rosvold et al., 1992, O'Dwyer, 1990, O'Dwyer et al., 1990). Both in vivo and in vitro studies indicated that, at concentrations below cytotoxicity, PALA potently inhibits CAD and reduces UTP and CTP pools. Nevertheless, even at the highest doses tested, *de novo* synthesis was not reduced below 25 % of control (Grem et al., 1988). Then, emphasis was shifted to clinical application of PALA in combination with other drugs, such as 5-fluorouracil (a pyrimidine analogue). However, the combinational drug trials showed that UTP pools were reduced up to 60% but *de novo* biosynthesis pathway could not be completely blocked (Grem et al., 1988). Although the mechanisms of clinical resistance to PALA were not clearly defined, several hypotheses have been proposed. The resistance could arise from the amplification of the cad gene and overproduction of the protein in sufficient amounts to overcome the inhibitory concentrations of PALA (Christopherson and Jones, 1980). It has also been shown that PALA could stimulate the utilization of CTP and UTP by the salvage pathway (Kensler et al., 1980). Another explanation is related with CAD functioning, since a blockage of the ATCase activity by PALA would cause a local increase in CP concentration that could compete with the inhibitor and overcome the inhibitory effect (Christopherson and Duggleby, 1983).

CAD and the other activities in *de novo* pyrimidine pathway do not only have the potential as antitumoral targets, but also in the treatment of malaria and other parasitic infections (Löffler et al., 2005). The malarian parasite *Plasmodium falciparum*, and other pathogens, exhibit rapid nucleic acid synthesis during their intraerythrocytic growth phase (Gero et al., 1984). These parasites do not have pyrimidine salvage pathways and depend exclusively on *de novo* pathway (Seymour et al., 1994). In *Plasmodium*, the first enzymes in the route are not fused together as in CAD. This is an important distinct feature that could be explored for novel drug development. So far, synthetic ribozymes with specificity for the *P. falciparum CPSase* gene have shown potent inhibitory effects on the growth of cultured *P.*

falciparum parasites with no toxicity to mammalian cells (Flores et al., 1997). Similarly, *de novo* pyrimidine synthesis is essential for the infective growth of the protozoan *Toxoplasma gondii.* Mutants lacking the first enzyme or the last enzyme in the *de novo* pyrimidine biosynthetic pathway are not pathogenic in mice and cannot replicate in culture in the absence of uracil (Fox and Bzik, 2002). The monofunctional CPSase and DHOase enzymes of *Toxoplasma* show distinctive properties compared with the activity of CAD in the mammalian host cell, thus becoming potential targets for therapeutic intervention (Asai et al., 1983).

The importance of *de novo* pyrimidines synthesis is also reflected in a number of syndromes and deficiencies associated with failures in distinct enzymes along the pathway (Ng et al., 2010, Suchi et al., 1997, Schwenger et al., 1993). However, until last year, there were no diseases directly attributed to a failure in CAD. Mutations in the *cad* gene in model organisms such as *Drosophila melanogaster* proved to be lethal when grown in pyrimidine-deficient media. When supplemented with a diet rich in uridine and uracil, the flies survived but suffered severe developmental problems and females were sterile (NØRby, 1970, Conner and Rawls Jr, 1982). Mutations in CAD in *Caenorhabditis elegans* also associated with a 95% mortality of late-stage embryos (Franks et al., 2006) and with abnormalities in the surviving 5% of the worms. Again, this lethality was overcome by supplying a media with a rich content in uridine and uracil. Similarly, the mutations in CAD in zebrafish (named *perplexed* mutations) were reported to cause abnormal development and embryonic lethality (Cox et al., 2014, Willer et al., 2005). Overall, these results indicated that the mutations compromising the activity of CAD were lethal, and provided an explanation to the fact that no human diseases had been associated with defects in CAD. However, the group led by Hudson Freeze recently reported the first partial CAD deficit in a 4-year old patient with severe problems in protein glycosylation (Ng et al., 2015). The patient presented an abnormal splicing resulting in a 63 amino acid deletion, within the CPSase domain in one allele and a missense mutation, R2024Q, in the ATCase domain in the other allele. The metabolic studies showed that these mutations impaired the incorporation of aspartate into RNA and DNA, and severely decreased the levels of UTP, CTP and all UDP-activated sugars for glycosylation.

5. The ATCase domain is key for the association of CAD into a 1.5 MDa particle

Despite the central role played in the metabolism of pyrimidines, the importance for cell proliferation and tumor development and the potential as a therapeutic target, by the time we started this study, there was a lack of detailed information about the architecture of CAD or about any of its functional domains. It was only known that this multifunctional polypeptide of ~240 kDa could self-assemble into ~1.5 MDa homo-hexamer, in equilibrium with other

oligomeric states such as trimers (Coleman et al., 1977, Lee et al., 1985, Qiu and Davidson, 2000).

The DHOase and ATCase domains of CAD from hamster have been isolated following limited proteolysis of the full-length protein (Davidson and Patterson, 1979, Davidson et al., 1981, Mally et al., 1981, Kelly et al., 1986, Huang et al., 1999) and have been obtained recombinantly from *E. coli* by expression of cDNAs of various lengths (Huang et al., 1999, Davidson et al., 1993, Davidson and Niswander, 1983, Maley and Davidson, 1988, Qiu and Davidson, 1998). It was not until 2014, when our group solved the crystal structure of the human DHOase domain that we had the first atomic glimpse of CAD (Grande-García et al., 2014, Lallous et al., 2012). This work also resulted in a revision that is presented as an appendix of this thesis (Ruiz-Ramos et al., 2015).

Many were the efforts to determine the crystal structure of the ATCase domain of CAD but with disappointing results (D. Evans personal communication). The group of Jeffrey Davidson proved that the recombinant ATCase domain of hamster CAD could complement the ATCase-deficient mammalian and bacteria cells, demonstrating that the ATCase domain of CAD is active when obtained separately from the other CAD domains (Davidson and Niswander, 1983, Major et al., 1989). It was also proven that the cleaved ATCase domain obtained by protease digestion or produced recombinantly formed homotrimers in solution (Grayson and Evans, 1983, Simmer et al., 1989, Qiu and Davidson, 1998), as described for all known bacterial ATCases (Shi et al., 2015). Davidson´s group also proved that the ATCase domain of CAD preserves specific interactions that were shown to stabilize the bacterial ATCase trimer (Qiu and Davidson, 1998). These conserved interactions were shown to be critical for the oligomerization of CAD since their disruption by site-directed mutagenesis (D90A or R269A) caused the dissociation of CAD hexameric particles (Qiu and Davidson, 2000). They also proved that the addition of the ATCase substrate, CP, or of the inhibitor PALA led to the re-association of CAD into hexamers (Qiu and Davidson, 2000). These results proved the key role of the ATCase domain for the assembly of CAD oligomers.

In the absence of a crystal structure of the ATCase domain of CAD, biochemical and modelling studies have supported the structural similarities with the *E. coli* ATCase catalytic subunits (Hemmens and Carrey, 1994, Scully and Evans, 1991, Maley and Davidson, 1988). The proposed similarity was further demonstrated by producing active chimeras that contained half of the ATCase domain of hamster CAD and half of *E. coli* ATCase catalytic subunit (Major et al., 1989).

6. *E. coli* ATCase as the paradigm of the transcarbamoylase superfamily

The transcarbamylase family groups a large number of enzymes catalyzing the transfer of the carbamyl group from carbamoyl phosphate (CP) to a nitrogen or oxygen atom in a receptor substrate. These enzymes are involved in important metabolic pathways, such

as pyrimidine and arginine biosynthesis (Labedan et al., 1999), agmatine fermentation (Polo et al., 2012) and allantoin degradation (Bojanowski et al., 1964), among others.

ATCase (EC 2.1.3.2) and ornithine transcarbamoylase (OTCase; EC 2.1.3.3) are the best-known members of this family. The ATCase from *E. coli* (ecATCase) was the first family member for which the structure was determined (Honzatko et al., 1982). Indeed, ecATCase is probably one of the best structurally characterized enzymes, with more than 60 different crystal structures determined. In addition to the exhaustive structural characterization, the ATCase from *E. coli* has been widely studied biochemically, becoming a textbook model to explain allosterism (Monod et al., 1965) and cooperativity mechanisms (Perutz, 1989).

The quaternary structure of the ecATCase is a dodecamer composed by two catalytic trimers bridged non-covalently face to face by three regulatory dimers (Gerhart and Schachman, 1965, Honzatko et al., 1982). The enzyme exists in two different structural and functional states: the low-activity, low-affinity T-state and the high-activity, high-affinity R-state (Figure 5A) (Changeux and Rubin, 1968). The conversation of the ecATCase holoenzyme from the T to R state occurs upon Asp binding to the holoenzyme in the presence of CP, The inhibitor PALA and ATP also trigger this T>R conversion (Gerhart and Pardee, 1962, Howlett and Schachman, 1977, Collins and Stark, 1971). Contrary, the CTP and UTP, the final product of *de novo* pathway, shift the equilibrium of the holoenzyme to the T-state (Gerhart and Pardee, 1962, Gouaux et al., 1990, Gerhart and Pardee, 1963). This T>R allosteric transition involves an elongation of the molecule of 11 Å along the three-fold axis, a relative rotation of 12º between the catalytic trimers and a 15º rotation of each regulatory dimer around their two-fold axis (Figure 5A) (Kantrowitz, 2012, Shi et al., 2015). The dissociation of the holoenzyme by heat treatment or with mercurial compounds results in more active catalytic trimers with non-cooperative active sites and no response to allosteric regulators. The isolated regulatory dimers binds nucleotide effectors but are catalytically inactive (Gibbons et al., 1974).

The ecATCase holoenzyme dissociation revealed that the functional unit of all transcarbamylases is a homotrimer with a concave triangular shape (Figure 5). Each subunit is divided into two similar domains. The N-terminal domains are involved in the binding of CP, therefore named N-domain or CP-domain, and are closer to the 3-fold axis of the trimer. The C-terminal domains protrude from the centre of the trimer and bind the second substrate. In the case of ATCase, this second substrate corresponds to aspartate (Asp) and is thus called Asp-domain. Both domains fold similarly, with a central parallel β-strand surrounded by α-helices connected by two additional transversal α-helices (Shi et al., 2015) (Figure 5). The active sites are located at the concave (bottom) face of the trimer, in a cleft formed between the two domains with residues from adjacent subunits (Jin et al., 1999) (Figure 5C). The binding of the substrates is ordered in all anabolic transcarbamoylases (Shi et al., 2015). CP binds in the first place to the N-domain, inducing the closure of a loop

(named loop 80s) from the adjacent subunit (Figure 5). Movement of the loop 80s interacts over the CP prepares the binding site for Asp (Wang et al., 2005). Upon Asp binding, a second loop from the C-domain (loop 240s) closes over the active site interacting with Asp and with the CP-domain (Figure 5) (Yuan et al., 1996). Releasing of the products is also order, delivering first the carbamoyl-aspartate (CAsp) and then, the inorganic phosphate.

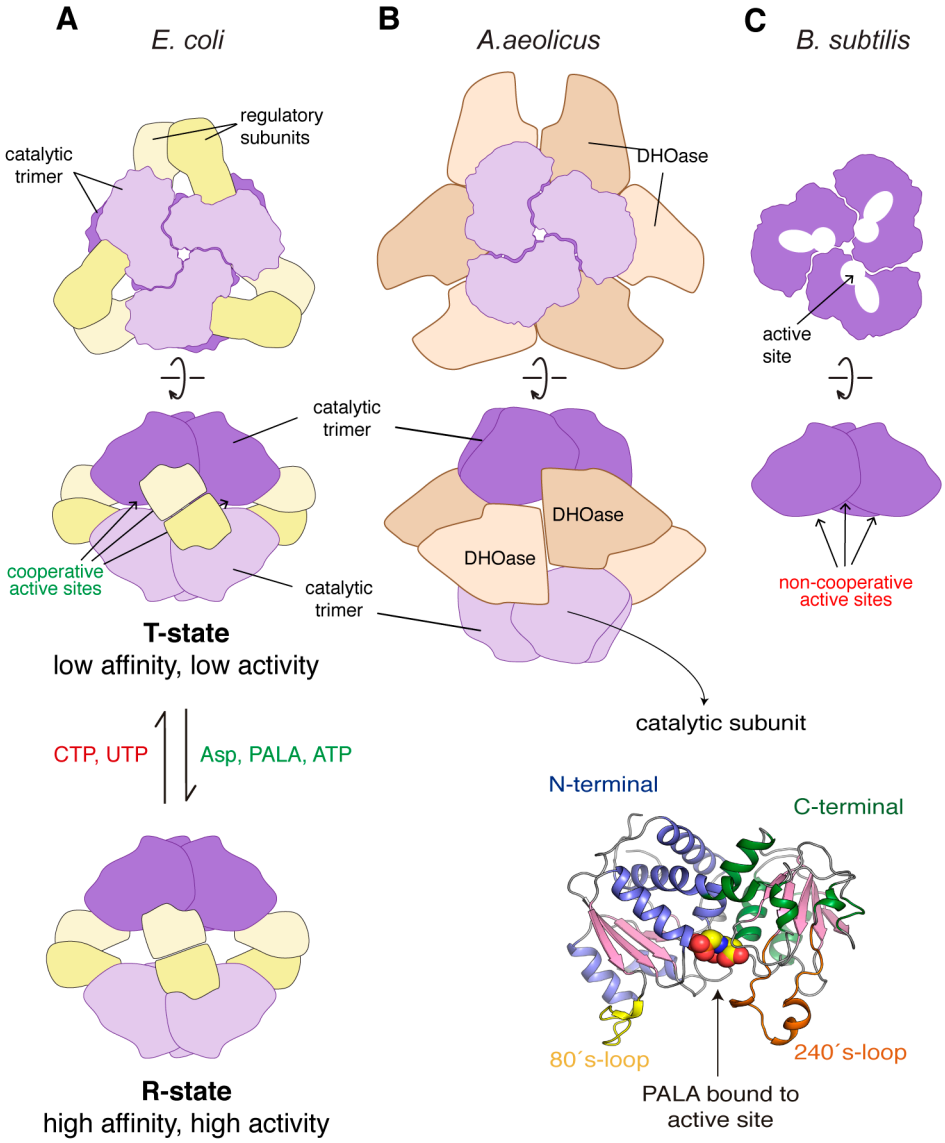

Figure 5. Schematic representation of the ATCases quaternary organization and allosteric transition in *E. coli* ATCase. (A) *E. coli* ATCase holoenzyme (Class B). The catalytic trimer is shown in purple, whereas the regulatory dimers are shown in yellow. The allosteric transition to the low and high activity conformations are shown below. **(B)** Representation of the *A. aeolicus* ATCase holoenzyme (Class A). The DHOase dimers are shown in brown. **(C)** Representation of the ATCase of *B. subtilis* (Class C). This simplest class does not interact with partners and lacks regulatory and allosteric properties. The crystal structure of one subunit of the *E. coli* ATCase trimer is shown bound to PALA in the active site.

There are numerous crystal structures of ecATCase, either for the catalytic trimers alone (ecATCase-C_3) or in complex with the regulatory subunits forming a holoenzyme (Kantrowitz, 2012, Kantrowitz and Lipscomb, 1990, Lipscomb and Kantrowitz, 2012, Schachman, 1987, Allewell, 1989).

The ATCases in prokaryotes and archaea can be classified into different groups depending on their quaternary structure and the interactions with other proteins (Bethell and Jones, 1969, Labedan et al., 2004) (Figure 5). ecATCase belongs to Class B, where two catalytic trimers associate through three regulatory dimers forming a dodecameric structure (Figure 5A). A particular case of Class B ATCase is that of *Thermotoga maritima*, where the regulatory and catalytic subunits are fused into a single polypeptide (Chen et al., 1998). A second type of ATCases (Class A) also exhibits a dodecameric arrangement, but the two catalytic trimers are connected by three dimers of DHOase. These DHOases can be active (e.g. *Aquifex aeolicus;* subclass A1) or inactive (e.g. *Pseudomonas aeruginosa;* subclass A2) (Figure 5B). In the case of *A. aeolicus*, it has been proven that the association with ATCase is necessary for the DHOase to be active (Ahuja et al., 2004). The third type (Class C) corresponds to the simplest form of ATCase, which is formed by isolated catalytic trimers with no interactions with any additional protein partners (e.g. *Bacillus subtilis*) (Stevens et al., 1991) (Figure 5C). Similarly to the ecATCase-C_3, the ATCases belonging to the Class C lack allosteric and cooperative mechanisms (Brabson and Switzer, 1975).

There are crystal structures available for ATCases belonging to either of the three classes in prokaryotes and archaea, with the current exception of subclass A2 (complex with inactive DHOases). However, we lack structural information about the eukaryotic ATCases. Despite the predicted similarity with prokaryotic counterparts (Scully and Evans, 1991), a detailed characterization of this domain in animals is important to understand the architecture, functioning and inhibition of CAD.

OBJECTIVES

The general aim of this thesis is to characterize structurally and biochemically the aspartate transcarbamoylase domain of human CAD (huATCase). This global goal implies attaining the following particular objectives:

1. Clone, express and purify the recombinant huATCase.
2. Crystallize huATCase free of ligands and in complex with the inhibitor PALA.
3. Determine the crystal structure of huATCase and to study at atomic level the interaction of the protein with the ligands.
4. Characterize the kinetic parameters of the isolated huATCase.

OBJETIVOS

El objetivo general de esta tesis es caracterizar de manera estructural y bioquímica el dominio aspartato transcarbamilasa de la proteína humana CAD (huATCasa). Para ello, se han de lograr los siguientes objetivos específicos:

1. Clonar, expresar y purificar la proteína recombinante huATCasa.
2. Cristalizar la huATCasa libre de ligandos y en complejo con el inhibidor PALA.
3. Determinar la estructura cristalográfica de huATCasa y estudiar a nivel atómico la interacción de la proteína con los ligandos.
4. Caracterizar los parámetros cinéticos del dominio aislado.

MATERIALS AND METHODS

1. Cloning of the ATCase domain of human CAD

1.1. Gene amplification and digestion of expression vectors.

The cDNA of human CAD (UniProt P27708) was purchased from Open Biosystems (clone ID 5551082). The gene fragment covering the ATCase domain of CAD, residues 1915–2225 (hereafter huATCase), was amplified by PCR using Deep Vent DNA Polymerase (New England Biolabs) and the oligos flanking-forward and flanking-reverse (Table 1). The PCR program consisted in an initial denaturalization step at 95 ºC for 2 min followed by 25 cycles of denaturalization (95 ºC for 30 s), annealing (58 ºC for 1 min) and extension (72 ºC for 1 min) with a final extension step of 10 min at 72 ºC. Flanking primers contained specific regions (Table 1, highlighted in bold) homologous to the 5' and 3' arms of the linearized pOPIN-F and pOPIN-M expression vectors (Oxford Protein Production Facility; https://www.oppf.rc-harwell.ac.uk/OPPF/). The expression vectors were linearized by digestion with the restriction enzymes KpnI and HindIII at 37 ºC for 3.5 h in buffer NEB 1 (Berrow et al., 2007). This buffer is optimized for KpnI, and thus, we added and excess of enzyme HindIII during the double digestion of the plasmids. The linearized vectors and the PCR amplified gene were loaded on 1 % agarose gels, and the corresponding bands were visualized by staining with Gel-Red (Biotium Inc.), excised and purified using QIAquick Gel Extraction kit (QIAGEN).

Table 1. Oligos used in huATCase cloning and site-directed mutagenesis. Adaptors for In-Fusion cloning are shown in italic and bold. The mutated triplets are shown in red.

Construct	Type	Sequence (5´→ 3´)
WT	Flanking-Forward	***AAGTTCTGTTTCAGGGCCCG***ATGTCACCCCTGCTGCACTCA
	Flanking-Reverse	***ATGGTCTAGAAAGCTTTA***GAAACGGCCCAGCACGGT
E1954A	Forward	ATGATGGTGCAGAAGGCGCGGAGCCTCGACATCC
	Reverse	GGATGTCGAGGCTCCGCGCCTTCTGCACCATCAT
D1958A	Forward	AAGGAGCGGAGCCTCGAGATCCTGAAGGGGAAGG
	Reverse	CCTTCCCCTTCAGGATCTCGAGGCTCCG CTCC
R2024Q	Forward	GTCGTCGTGCTCCAGCACCCCCAGCCTGG
	Reverse	CCAGGCTGGGGGTGCTGGAGCACGACGAC

1.2. In-Fusion cloning.

All the gene constructs in this work were made by ligation-independent cloning using In-Fusion HD cloning kit (Clontech Laboratories) (Figure 6). This technology allows a fast and directed insertion of a fragment of DNA in any vector by recombination of 15 bp overlapping regions (Table 1, highlighted in bold). The pOPIN vector suite is designed for the insertion of a single PCR amplicon flanked by the proper adaptors into a number of vectors that add different fusion tags to either the N- or C-termini of the protein. In particular, the chosen pOPIN-F and pOPIN-M vectors introduce a His_6-tag or a His_6-maltose binding protein (MBP) tag, respectively, at the N-terminus of huATCase. This technology, which is not limited by the presence of undesired restriction sites within the gene of interest, is particularly

convenient for cloning of long DNA fragments such as full-length CAD or the bi-functional DHOase-ATCase construct (see materials and methods, section 10.1.).

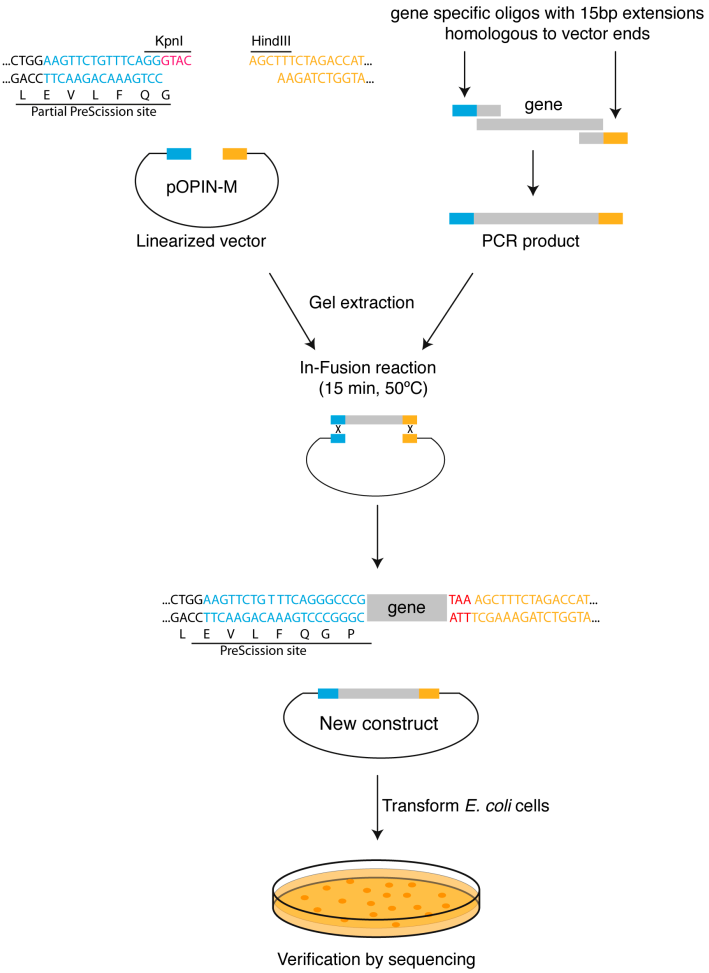

Figure 6. Scheme of the In-Fusion cloning protocol. The expression vector is linearized by KpnI and HindIII restriction enzymes. The N-terminal sequence of the vector corresponds to the PreScission cleavage site (in blue). The vector extremes are sequence complementary to the ends of the PCR product, which are inserted by engineering the flanking oligos. Both, vector and PCR product, are mixed in the In-Fusion reaction, where the recombinase inserts the PCR product inside the vector. The reaction is directly used for transformation of *E. coli* TOP10 cells.

In general, the In-Fusion reaction was carried out at a 3-fold excess molar ratio of insert and vector, fixing the vector amount to a minimum of 100 ng per reaction. The mixture was incubated for 15 min at 50 ºC and 2.5 µl were transformed straight away in chemo-competent *E. coli* TOP10 cells by heat-shock (Inoue et al., 1990). A quick screen of bacteria colonies containing the plasmid was carried by PCR. For this, a fraction of the grown colonies was picked with a pipette tip and transferred to a PCR tube with 12.5 µl of sterile water, which was subsequently heated at 95 ºC for 10 min. Then, 12.5 µl of a 2X PCR mixture containing a 5' primer for the vector and a 3' primer for the gene (flanking-reverse) was added to the sample and the PCR was performed as previously described. The positives colonies were detected by the visualization of an amplified band of the expected size in a 1 % agarose gel. The positive colonies were re-picked and grown in media for plasmid purification (mini-prep). The correct cloning of huATCase gene within pOPIN-F (F-huATC) and pOPIN-M (M-huATC) was verified by sequencing of the complete gene and the corresponding flanking sequences.

1.3. Site-directed mutagenesis

huATCase single-point mutations E1954A, D1958A and R2024Q were made by directed-mutagenesis using the "overlapping extension PCR" method (Ho et al., 1989) (Figure 7). Firstly, two independent PCR reactions (PCR I and PCR II) were performed using a flanking primer (flanking-forward or flanking-reverse) and internal primers with the desired mutations (Table 1). The PCR program consisted in an initial denaturalization step at 98 ºC for 5 min followed by 30 cycles of denaturalization (95 ºC, 30 s), annealing (58 ºC, 45 s) and extension (72 ºC, 1 min) with a final extension step at 72 ºC for 5min. The PCR products were purified from a 1 % agarose gel using the QIAquick Gel Extraction Kit (QIAGEN) and applied as templates for a third PCR reaction (PCR III) with the primers flanking-forward and flanking-reverse. The products resulting from this PCR III were gel-purified and cloned in the linearized pOPIN-M using In-Fusion. The presence of the mutation and the absence of other undesired mutations were verified by gene sequencing.

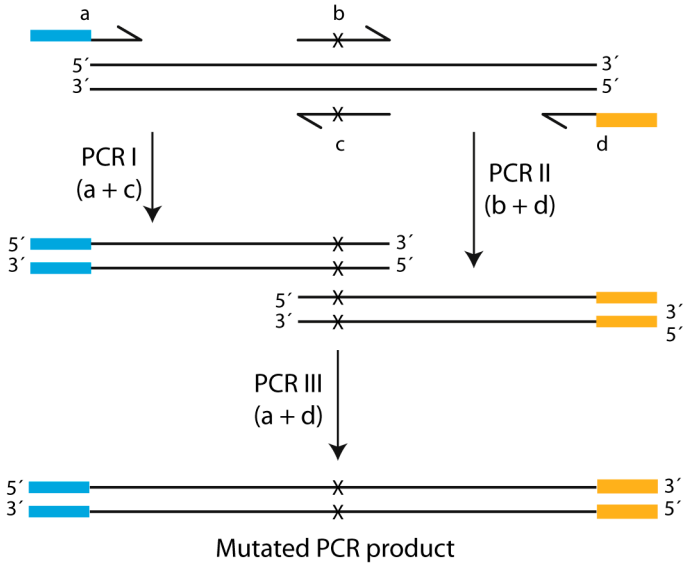

Figure 7. Site-directed mutagenesis protocol. This protocol involves two separated PCR reactions (PCR I and PCR II) to amplify the gene from the ends to the mutation region. The crosses indicate the point mutation. The PCR fragments are gel-purified and used as template for a third PCR reaction (PCR III) using both flanking oligos. The mutated PCR product contains the extensions for In-Fusion cloning (shown in blue and yellow).

2. Production of the ATCase domain of human CAD

2.1. Protein expression

E. coli BL21 (DE3) pLysS cells (Novagen) transformed with F-huATCase and M-huATCase plasmids were grown in Luria-Bertani (LB) liquid media, supplemented with 100 mg ml^{-1} ampicillin, at 37 ºC in a shaking incubator (220 rpm) to a density of OD = 0.6 - 0.8. Then, protein expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and cultures were grown for six hours at 37 ºC before to harvesting by centrifugation. Alternatively, prior to the addition of the IPTG, the cultures were chilled in an ice-water bath for 10 min, and expression was performed overnight at 18-20 ºC.

Protein expression was also carried out by auto-induction in high density cultures (Studier, 2005). Briefly, a transformed colony of *E. coli* BL21 (DE3) pLysS cell was grown in

50 ml ZY-0.8G (Table 2) medium supplemented with 100 mg ml^{-1} ampicillin and 34 mg ml^{-1} chloramphenicol for 16 hours at 37 ºC in a shaking incubator (220 rpm). Next day, 1 L of ZY-5052 medium (Table 2), supplemented with the same antibiotics, was inoculated with 20 ml of the overnight pre-culture and grown for 6 hours at 37 ºC at 220 rpm. Then, temperature was decreased to 20 ºC for 20 h. Cells were harvested by centrifugation in a Beckman Coulter Avanti J-20 XP centrifuge using 1 L polypropylene bottles, washed with phosphate-buffered saline (PBS), flash-frozen and stored at -80ºC.

Table 2. Autoinduction media components.

Stock solutions	ZY-0.8G Final concentration	ZY-5052 Final concentration
ZY[a]	To make up volume	To make up volume
1M $MgSO_4$	1 mM	1 mM
20% glucose	0.8 %	-
50x 5052[b]	-	1x
20x NPS[c]	-	1x
Ampicillin	100 µg/ml	100 µg/ml
Chloramphenicol	34 µg/ml	34 µg/ml

[a] ZY medium (10 g bactotryptone, 5 g yeast extract in 1 L distilled water).
[b] 50x 5052 solution (25 % glycerol, 2.5 % glucose, 10 % α-lactose).
[c] 20x NPS solution [1 M Na_2HPO_4, 1 M KH_2PO_4, 0.5 M $(NH_4)_2SO_4$].

2.2. Protein purification

The cells were thawed and resuspended in chilled buffer A (20 mM Tris–HCl pH 8, 0.5 M NaCl, 10 mM imidazole, 5% glycerol, 2 mM β-mercaptoethanol; 10 ml of buffer per mg of cell paste) complemented with 2 mM phenylmethanesulfonylfluoride (PMSF) and disrupted by sonication using a Vibra-Cell 75042 sonicator (Bioblock Scientific). The lysate was clarified by centrifugation at 40,000 rpm for 40 min at 4 ºC using a Beckman Coulter Optima XL-100K ultracentrifuge and a Beckman 45 Ti rotor. The supernatant was passed through a 0.45 µm pore filter and applied onto a 5 ml Ni^{2+}-loaded HisTrap Chelating FF column (GE Healthcare, USA) using a peristaltic pump at a flow-rate of 1 ml min^{-1}. After loading, the column was connected to an ÄKTA Prime or a FPLC Purifier (both from GE-Healthcare). Following extensive washing of the column with 20 column volumes of buffer A containing 34 mM imidazole, the protein was eluted step-wise with buffer A supplemented with 250 mM imidazole. The sample was diluted to 1.3 mg ml^{-1} with buffer S (20 mM Tris–HCl pH 7, 75 mM NaCl, 5% glycerol, 2 mM β-mercaptoethanol) to prevent protein precipitation and dialyzed overnight against the same buffer. During this step, the fusion tag was cleaved off by inclusion of Prescission protease ($1/20^{th}$ of the protein weight) in the dialysis bag. The dialyzed and cleaved sample was loaded on a 5 ml HiTrap SP HP column connected to a 5 ml GSTrap FF column (both from GE Healthcare, USA), both pre-equilibrated in buffer S. huATCase was retained in the HiTrap SP column, whereas the GST-tagged PreScission protease was retained in the GSTrap column and both the His_6-MBP tag and the undigested huATCase passed through both columns. Then, huATCase was eluted at 110 mM NaCl in a

linear gradient of salt. The fractions containing the protein were pooled and loaded directly on a 5 ml HiTrap Heparin column (GE Healthcare, USA) pre-equilibrated in buffer S. Then, huATCase eluted in a NaCl gradient at 330 mM. The protein was concentrated to 3-5 mg ml^{-1} in an Amicon Ultra system with a 30-kDa cutoff membrane. At this step, the protein could be conserved by adding 30% dimethyl sulfoxide (DMSO) and 5% glycerol, flash-frozen in liquid nitrogen and stored at -80 ºC. As a last purification step, the sample was further purified by size-exclusion chromatography on a Superdex 200 10/300 column (GE Healthcare, USA) pre-equilibrated in buffer GF [20 mM Tris- acetate pH 8.3, 0.15 M NaCl, 2% glycerol, 0.2 mM tris(2-carboxyethyl)phosphine (TCEP)]. huATCase eluted as a single peak that was concentrated to 3 mg ml^{-1} by ultracentrifugation. All purification steps were carried out at 4 ºC to preserve protein activity and to prevent degradation. Sample purity was evaluated by SDS–PAGE and Coomassie staining. Since huATCase has a low theoretical extinction coefficient (11920 M^{-1} cm^{-1}) due to the lack of tryptophan residues, protein concentration was measured by Bradford method (Bradford, 1976) using the Bio-Rad Protein Assay Dye.

huATCase-D1958A mutant was purified as the wild-type (WT) protein. In turn, mutant E1954A eluted from the S-column in a broad peak along the salt gradient and thus, to avoid excessive protein dilution, the elution was done step-wise with 500 mM of salt. The fractions containing the huATCase-E1954A were pooled and diluted with dilution buffer (20 mM Tris-HCl pH7, 5 % glycerol and 2 mM β-ME) to a final NaCl concentration of 250 mM. Then, the sample was loaded into the heparin column and followed the regular protocol. On the other hand, the solubility of mutant R2024Q in buffer S containing 75 mM NaCl was low, and thus, a second Ni^{2+}-affinity column replaced the ion-exchange chromatography step. huATCase-R2024Q eluted from the first Ni^{2+}-affinity column was digested and dialysed against buffer A. Then, the sample was passed through a second Ni^{2+}-affinity column, equilibrated in buffer A, from which the digested huATCase-R2024Q eluted by increasing the imidazole concentration to 34 mM. Then, the sample was diluted with dilution buffer to a final concentration of 150 mM NaCl and applied to the heparin column. R2024Q mutant eluted from the heparin column at 220 mM NaCl and purification continued as for the WT.

2.3. Production of GST- PreScission protease

The pGEX-4-T plasmid containing the PreScission protease gene –also called HRV 3C protease– was a kind gift of John Kuriyan (University of California, Berkeley, USA). This plasmid allows the expression of the protease N-terminal tagged to a glutathione S-transferase (GST) removable by thrombin cleavage. The plasmid was transformed into *E. coli* BL21 (DE3) pLysS cells. A 50 ml LB pre-culture supplemented with 100 mg ml^{-1} ampicillin and 34 mg ml^{-1} chloramphenicol was inoculated with one colony and grown overnight at 37 ºC. 20 mL of pre-culture were inoculated in 1 L of LB medium supplemented with the antibiotics and grown at 37 ºC until reaching a OD_{600}= 1.0. Then, the culture was

quickly chilled for 10 min in an ice-water bath. The culture was grown for an additional hour at 20 ºC and then, protein expression was induced with 0.5 mM IPTG for 16 h at 20 ºC. Cells were harvested by centrifugation, washed with PBS and stored at -80 ºC.

For protein purification, the cell paste was resuspended in lysis buffer (1X PBS, 1 mM DTT and 5 mM EDTA; 5 ml per mg of cell paste; without adding any protease inhibitor). Cells were lysed by sonication and clarified by centrifugation at 40,000 rpm for 30 minutes. The supernatant was filtered and loaded twice in a 5 ml GSTrap FF column (GE Healthcare, USA) at 1 ml min^{-1} using a Peristaltic pump. The column was washed with 200 ml of lysis buffer, and the GST-tagged PreScission protease was eluted with 10 ml of 20 mM reduced-glutathione pH 7.2. The eluted protein was dialyzed against buffer Q (20 mM Tris-HCl pH 8, 150 mM NaCl and 1mM DTT) and loaded into a 5 ml HiTrap Q HP column (GE Healthcare, USA) pre-equilibrated in the same buffer. The protease did not bind to the column, and was concentrated to ~1.5 mg ml^{-1} using an Amicon with a 30-KDa cutoff membrane. Then, the concentrated protein was supplemented with 20 % glycerol, flash-frozen in liquid nitrogen and stored at -80ºC ready to use. All purification steps were carried out at 4 ºC.

3. Size-exclusion chromatography coupled to multi-angle light-scattering (SEC-MALS) measurements

For molar-mass determination, we used 400 µl of purified huATCase wild-type (huATCase-WT) and of E1954A and D1958A mutants at 1.4 mg ml^{-1}. Proteins were fractionated by gel filtration on a Superdex 200 10/300 column pre-equilibrated with buffer GF using an ÄKTA purifier at a flow-rate of 0.5 ml min^{-1}. For huATCase-WT and huATCase-D1958A, fractions corresponding to the first half of the peak (~2 ml) were pooled and re-injected at a concentration of 0.25 mg ml^{-1}, whereas mutant E1954A was re-injected at 0.16 mg ml^{-1}. The mutant huATCase-R2024Q was injected at 1.1 mg ml^{-1} and no re-injection was performed. In all cases, the eluted sample was characterized by in-line measurement of the refractive index and multi-angle light scattering using Optilab T-rEX and DAWN 8+ instruments (Wyatt Technology), respectively. Data analysis was performed using the ASTRA 6 software (Wyatt Technology) to obtain the molar mass (Wyatt, 1993). The program *GraphPad Prism* was used for graphical representation.

4. Sedimentation velocity measurements

Analytical ultracentrifugation (AUC) studies were performed by J. R. Luque-Ortega at the *Unidad de Ultracentrifugación Analítica y Dispersión del Luz* - Centro de Investigaciones Biológicas (CIB-CSIC). Sedimentation velocity runs were carried out at 48,000 rpm and 20 ºC in an XL-I analytical ultracentrifuge (Beckman-Coulter, Inc.) equipped with UV-visible and Raleigh interference detection systems, using an An-50 Ti rotor and 12-mm double sector centerpieces. The experiments were performed with huATCase-WT, huATCase-E1954A and

huATCase-D1958A samples. Prior to measurements, protein buffer was exchanged to 20 mM Tris-HCl pH 8.3, 0.15 M NaCl, 1 % glycerol and 0.1 mM TCEP by dialysis. The concentrations of huATCase-WT tested were 1.4, 0.4, 0.1, 0.06 and 0.03 $mg \cdot ml^{-1}$ (40.2, 11.5, 2.9, 1.7 and 0.1 µM). The concentrations of huATCase-E1954A were 1.0, 0.5, 0.38, 0.25, 0.13, 0.06 and 0.03 $mg \cdot ml^{-1}$ (28.7, 14.3, 10.9, 7.2, 3.7, 1.8 and 0.9 µM). For huATCase-D1958A the concentrations were 2, 1, 0.5, 0.25, 0.125, 0.06, 0.03 $mg \cdot ml^{-1}$ (57.4, 28.7, 14.3, 7.2, 3.6, 1.8 and 0.9 µM). Sedimentation profiles were registered by absorbance at 230 nm. The sedimentation coefficient distributions were calculated by least squares boundary modelling of sedimentation velocity data using *c*(*s*) method as implemented in SEDFIT (Schuck, 2000). Measurements were normalized to obtain the percentage of trimers and monomers at each protein concentration. The program *GraphPad Prism* was used for graphical representation.

5. Thermal shift assays

Thermal stability experiments were carried out using a 7300 Real-Time PCR System (Applied Biosystems). Two types of experiments were performed to characterize the huATCase-WT and mutants: pH buffer screening and substrate-induced stability. For pH screening, the pH Buffer Screen (Emerald Bio) was used as an additive screen. It contains 12 different buffer systems at 8 different pH points, from 2.4 to 11.6. Each reaction contained 2 µM huATCase protein, 200 mM NaCl, 2% glycerol, 2.5x SYPRO Orange (Invitrogen) and 50 mM pH additive buffer in a final volume of 40 µl. The samples were dispensed in a MicroAmp optical 96-well reaction plate (Applied Biosystems) and sealed with an optical adhesive film (Applied Biosystems). Fluorescence intensity changes were monitored by increasing the temperature ramp by 1 ºC per minute from 20 ºC to 95 ºC.

Ligand-induced stability was tested in 40 µl reactions containing 1 µM huATC, 20 mM Tris-acetate pH 8.3, 100 mM NaCl, 2% glycerol, 2.5x SYPRO Orange and CP or PALA at concentrations ranging from 4 µM to 1 mM. Asp and succinate (Suc) were tested up to 30 mM. PALA (CAS 51231-79-0) was obtained from the Developmental Therapeutics Program (NCI, National Institutes of Health) Open Chemical Repository. Data were processed and analysed with an Excel script (ftp://ftp.sgc.ox.ac.uk/pub/biophysics) (Niesen et al., 2007) and plotted with *GraphPad Prism* software.

6. Enzymatic assays

6.1. Substrate saturation assays

huATCase activity was assayed by a colorimetric method for carbamoyl-aspartate (CAsp) quantification (Prescott and Jones, 1969) adapted to a 96-well plate format (Else and Hervé, 1990). The materials used for the reaction included:

- Carbamoyl phosphate (CP): the CP stocks were prepared at 50 mM by dissolving lithium carbamoyl phosphate dibasic hydrate (Sigma-Aldrich) in chilled water and flash-frozen immediately. CP aliquots were stored at -20 ºC and thawed on ice for single use right before performing the experiments.
- L-Aspartic acid (L-Asp): Asp (Sigma Aldrich) was dissolved in water at 300 mM final concentration. pH was adjusted to 8.3 with NaOH. L-Asp solution can be stored at 4 ºC for one week.
- Carbamoyl-L-aspartate (L-CAsp): L-CAsp (Bachem) was dissolved in water at 200 mM final concentration and stored at -20 ºC. To perform the standard curve, L-CAsp was diluted at 2mM with 50 mM Tris-acetate pH 8.3 buffer.
- N-(phosphonacetyl)-L-Aspartate (PALA): PALA (CAS 51321-79-0) was obtained from the Developmental Therapeutics Program (NCI, National Institutes of Health) Open Chemical Repository. The PALA stock was prepared at 100 mM in water and pH was adjusted to 7 with NaOH. PALA aliquots were stored at -20 ºC.
- Succinate (Suc): Succinate (Sigma Aldrich) was dissolved in water at 0.5 M final concentration. pH was adjusted to 8.4 with NaOH. Suc stock was stored at 4 ºC for one week.
- Colour mix solution: Colour mix consist of two parts of reagent A (0.37% antipyrine and 0.25% ammonium iron(III) sulfate in 25% H_2SO_4 95% and 25% H_3PO_4 85%) and one part of reagent B (0.4% diacetylmonoxime in 7.5% NaCl) (Nuzum and Snodgrass, 1976). The two reagents were gently mixed right before usage, and stored at 4 ºC in the dark until added to the reaction.
- Protein samples: reactions were performed with 0.04 µM huATCase WT, 0.04µM D1958A mutant, 0.08 µM E1954A mutant and 0.4 µM R2024Q mutant.

Reactions were carried out at 25 ºC in 50 mM Tris-acetate pH 8.3 and 0.1 mg ml^{-1} bovine serum albumin (BSA) in a final volume of 150 µl (Stebbins and Kantrowitz, 1989). In general, huATCase was pre-incubated with Asp for 10 min in a 96-well plate floating in a water bath at 25 ºC. The reaction was then triggered by adding CP. Substrate saturation experiments were done varying one substrate concentration while keeping the other fixed (10 mM Asp or 5 mM CP). Asp concentration varied between 0 and 30 mM, while CP varied from 0 to 10 mM. Reactions were stopped after 5 min by addition of 100 µl of colour mix solution (Else and Hervé, 1990). The plate was sealed with adhesive plastic and boiled at 95 ºC for 15 min to allow colour developing (LM Polo, personal communication). Then, the plate was cooled down in the dark for 30 min. After a short spinning of the plate, absorbance was measured at 450 nm in a Victor 1420 Multilabel Counter (Perkin Elmer). A standard colour curve was performed with serial dilutions of L-CAsp in 150 µl of final volume in the same 96-well plate as the titration assays. The process of colour developing and measurements was

carried out as already explained for titration assays. Data analysis was performed with *GraphPad Prism* software. huATCase activity was defined as mmol of CAsp synthetized per hour per milligram of protein.

Activity assays were also carried out in presence of the inhibitor PALA. In this case, the protein was pre-incubated with increasing concentrations of PALA (0 to 0.2 μM) and Asp for 10 min at 25 ºC. Reaction was triggered by adding CP. Asp and CP concentrations were the same used in substrate saturation experiments. Reactions were stopped after 5 min with colour mix solution. To test whether PALA is a slow-tight binding inhibitor, activity assays were performed at increasing stoichiometric concentrations of huATCase and PALA, ranging from 0 to 0.4 μM. The reaction was stopped at different times, up to 10 min, and colour was developed and measured as explained. Additionally, activity assays were performed at PALA concentrations up to 100 μM and at fixed 10 mM Asp concentration. Reaction was triggered with 5 mM CP and stopped after 5 min with colour mix reaction.

Alternatively, to substrate saturation experiments, huATCase was pre-incubated with PALA and varying CP concentrations for 30 min at 4 ºC. After the incubation for 10 min at 25 ºC, the reaction was triggered by adding 10 mM Asp and stopped after 5 min by adding 100 μl of colour mix solution. The colour was developed and measured as explained.

Succinate activation assays were performed by pre-incubating the protein with Asp at sub-saturating concentration (0.2, 0.5 and 1 mM) and succinate (from 0 to 25 mM) for 10 min at 25 ºC. The reaction was triggered with 5 mM CP and stopped after 5 min with colour mix solution. Colour was developed and measured as explained.

7. Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) studies were performed by Adrián Velazquez-Campoy at the Institute of Biocomputational and Physics of Complex Systems (BIFI). Experiments were performed in an Auto-ITC200 calorimeter (MicroCal, GE Healthcare) at 15 ºC. Calibration and performance tests of the calorimeter were carried out conducting Ca^{2+}-EDTA titrations with solutions provided by the manufacturer. The reaction cell contained 200 μl of protein at 30- 40 μM in buffer GF. Titration experiments consisted on 19 injections of 2 μl of PALA or CP at 0.4 mM and 1 mM, respectively. To measure CP binding in presence of succinate (Suc), the protein was pre-mixed with 2 mM Suc in the cell. All solutions were prepared in buffer GF and degassed before titrations. Titrant was injected at appropriate time intervals to ensure that the thermal power signal returned to the baseline prior to the injection. To achieve homogeneous mixing in the cell, the stirring speed was kept constant at 1,000 rpm. The results were measured by the heat per injection normalized per mol of injectant *versus* molar ratio. The data were analysed with Origin 7 (OriginLab) using one-site and two-site binding models (Moreno-Beltrán et al., 2015).

8. Crystallization

8.1. Optimization of huATCase crystals

High-throughput crystallization trials of huATCase were performed using a Cartesian MicroSys robot (Genomic Solutions), by the sitting-drop vapour-diffusion method in 96-well MRC plates. Nanodrops consisting of 0.2 µl of freshly purified protein plus 0.2 µl of reservoir solution were equilibrated against 60 µl of reservoir solution. Preliminary screening involved different protein concentrations and the following commercial crystallization screens: The JCSG+, PACT, Protein Complex, pHClear and pHClear II Suites (all from Qiagen). Plates were placed in a Rock Imager at 18 ºC and crystal growth was monitored using the Rock Maker software (Formulatrix). Initial hits were optimized manually varying the protein and precipitant concentrations, as well as the pH, in 15-well plates (Qiagen) with hanging drops consisting of 1 µl of protein solution and 1 µl of reservoir solution. Optimized crystallization condition for huATCase free of ligands consisted of protein at 1.5 mg ml^{-1} in GF buffer and 0.1 M Tris pH 8.5, 6% PEG 8000 and 9% ethylene glycol (EG) as mother liquor at 18 ºC. Free-ligand crystals were cryo-protected by gradual increasing the EG up to 30% and frozen in liquid nitrogen for subsequent diffraction studies.

8.2. Co-crystallization of huATCase and PALA

To crystallize the protein with the inhibitor PALA, huATCase at 3 mg ml^{-1} in buffer GF was diluted with an equal volume of 0.1 M CAPS pH 10 and 0.15 M NaCl to change pH and prevent protein precipitation upon PALA binding. PALA was added to a final concentration of 0.5 mM and incubated on ice for ~30 min. Then, the sample was centrifuged for 10 min at 13,000 rpm and used for crystallization. Optimal condition for the co-crystallization with PALA was obtained using 10 mM $ZnCl_2$, 15% PEG 6000 and 50 mM sodium acetate pH 4.8 as mother liquor at 18 ºC. Prior to cooling in liquid nitrogen, the PALA crystals were transferred to mother liquor supplemented with 20% glycerol for cryo-protection.

8.3. Soaking of huATCase crystals with CP

huATCase crystals free of ligands grown at a protein concentration of 1.73 mg ml^{-1} in similar conditions as described above (0.1 M Tris pH 8.5, 6% PEG 8000, 11% EG) and were soaked in 50 µl of mother solution supplemented with 5 mM CP for 5 min. Crystals were quickly cryoprotected in the same solution supplemented with 30 % EG and flash-frozen in liquid nitrogen.

9. Data collection and structure determination.

X-ray diffraction data were collected at 100 K on beamlines PX-I (SLS synchrotron, Villigen) and XALOC (ALBA synchrotron, Barcelona) using Pilatus 6M detectors. Data processing and scaling were performed with XDS (Kabsch, 2010). Analysis of the intensity

distribution and statistics were performed with phenix.xtriage from the PHENIX package (Adams et al., 2010). Crystallographic phases were determined by molecular replacement using PHASER (McCoy et al., 2007). The model was built with Coot (Emsley et al., 2010) and refined with PHENIX (Adams et al., 2010). Validation of the model was determined with MolProbity (Chen et al., 2010).

10. Production of the bifunctional human DHOase-ATCase construct

10.1. Cloning and site-directed mutagenesis

The gene fragment covering the DHOase and ATCase domains of human CAD, and the 91 aminoacids inter-domain linker (huDHOaseATCase; residues 1456–2225) was amplified by PCR using the oligos flanking-forward and flanking-reverse (Table 3) and cloned in pOPIN-F using In-Fusion as previously described (section 1.1). Giving the large size of the insert (2,310 bp), the PCR extension time was lengthened to 2 min, with a final extension step of 10 min. Site-directed mutagenesis was carried out also as described in section 1.3. The pairs of oligonucleotides used to introduce the single-point mutation M1601E, D2009A, S1859D and S1859E are shown in Table 3. Wild-type and the mutated variants were verified by complete sequencing of the gene and flanking vector regions.

Table 3. Oligos used in huDHOaseATCase cloning and site-directed mutagenesis. The adaptors for In-Fusion cloning are shown in italic and bold. The mutated triplets are shown in red.

Construct	Type	Sequence (5´→ 3´)
WT	Flanking-Forward	***AAGTTCTGTTTCAGGGCCCG***ATGACCTCCCAAAAGCTTGTG
	Flanking-Reverse	***ATGGTCTAGAAAGCTTTA***GAAACGGCCCAGCACGGT
M1601E	Forward	GCTGCTGTCCTCGAGGTGGCTCAGCTC
	Reverse	GAGCTGAGCCACCTCGAGGACAGCAGC
D2009A	Forward	GGCGAATCCCTGGCTGCCTCCGTGCAGACCATG
	Reverse	CATGGTCTGCACGGAGGCAGCCAGGGATTCGCC
S1859D	Forward	CGAATCCATCGAGCCGATGACCCAGGTTTGCCAGC
	Reverse	GCTGGCAAACCTGGGTCATCGGCTCGATGGATTCG
S1859E	Forward	CGAATCCATCGAGCCGAAGACCCAGGTTTGCCAGC
	Reverse	GCTGGCAAACCTGGGTCTTCGGCTCGATGGATTCG

10.2. Protein expression in HEK293 cells

huDHOaseATCase wild-type (huDHOaseATCase-WT) and the mutants were expressed in HEK293S GnTI$^-$ cells grown in 1 L suspension cultures in FreeStyle medium (Invitrogen) supplemented with 1% fetal bovine serum (Aricescu et al., 2006) at 37 ºC and 5% CO_2 in an orbital shaker at 135 rpm. Cultures were divided into two baffled polycarbonate Erlenmeyer flasks (CORNING) with 2 L capacity for better aeration. Cells were transfected at a density of 1.5 million cells ml^{-1} by adding 50 ml of a 1:3 ratio mixture of DNA plasmid and polyethylenimine (PEI 25 kDa branched; Sigma). Prior to the transfection, the DNA and PEI were diluted to 20 and 60 mg ml^{-1}, respectively, in UltraDOMA medium (Lonza) and incubated for 5 minutes at room temperature. Then, the solutions were mixed together and

incubated for 10 min at room temperature before adding the mixture to the cells. After the transfection, the cells were grown for 2–3 days and then, harvested by centrifugation. The cell pellet was washed with PBS and store at -80ºC.

10.3. Protein purification

The cells were resuspended in 50 ml of chilled buffer A supplemented with 3 mM β-ME, 2 mM PMSF, 1 µM leupeptin and 1 µM pepstain A. Cells were disrupted in a chilled glass Dounce tissue grinder, followed by 3 min sonication (pulses of 3 s and 1 s breaks). The clarified supernatant was loaded onto a 5 ml Ni^{2+} HisTrap Chelating FF column (GE Healthcare, USA) using a peristaltic pump. Following extensive column washing of 30 column volumes with buffer A containing 60 mM imidazole, the protein was eluted with buffer supplemented with 300 mM imidazole. Fractions containing the eluted protein were pooled and dialysed overnight against buffer D (20 mM Tris-HCl pH8, 0.5 M NaCl, 5% glycerol, 40 mM imidazole and 2 mM β-ME). PreScission protease was added in the dialysis bag to cleave off the His_6-tag. After dialysis, the protein was centrifuged at 4,000 rpm for 10 min to eliminate the precipitate, filtered and loaded into a 5 ml Ni^{2+} HisTrap Chelating FF column attached to a GSTrap FF column (GE Healthcare, USA) and pre-equilibrated in buffer D. The digested huDHOaseATCase protein did not bind to these columns, and was pooled and concentrated up to 1.5 mg ml^{-1} in an Amicon Ultrasystem with a 30K cutoff membrane. As a final step, the protein was loaded into a Superdex 200 10/300 pre-equilibrated in buffer GF2 (20 mM Tris-HCl pH 8, 0.3 M NaCl, 5 % glycerol and 0.1 mM TCEP). The elution peak was concentrated as before, supplemented with 20% glycerol, flash-frozen in liquid nitrogen and stored at -80 ºC.

11. SEC-MALS analysis

Molar-mass determination was carried out as previously described in section 3. 400 µl of purified protein at approximately 1 mg ml^{-1} were fractionated in a Superose 6 10/300 column equilibrated in buffer GF2 using an ÄKTA purifier at a flow-rate of 0.5 ml min^{-1}. huDHOaseATCase-WT was first injected at 1.72 mg ml^{-1} and the 250 µl fraction corresponding to the maximum of the peak (fraction E2) was re-injected at a concentration of 0.32 mg ml^{-1}. The mutants M1601E, D2009A, S1859D and S1859E were applied to the column at 0.9, 1.0, 1.2 mg ml^{-1} and 1.1 mg ml^{-1}, respectively. Mutants were not re-injected in the column. The eluted samples were characterized by in-line measurement of the refractive index and multi-angle light scattering using Optilab T-rEX and DAWN 8+ instruments (Wyatt Technology), respectively. Data analysis was performed using the ASTRA 6 software (Wyatt Technology) to obtain the molar mass (Wyatt, 1993). The program *GraphPad Prism* was used for graphical representation.

RESULTS

1. The ATCase domain of human CAD can be produced recombinantly

1.1. Cloning of human ATCase domain

Based on sequence comparison with bacterial homologues and on previous work on hamster CAD by the group of Jeffrey Davidson (Maley and Davidson, 1988, Major et al., 1989), we mapped the ATCase domain within human CAD, starting in residue S1915 and ending at F2225, the most C-terminal CAD residue (Figure 8). The corresponding region of the human *cad* gene (nucleotides 5768 to 6704) was amplified by PCR and inserted in two expression vectors from the pOPIN series (Oxford Protein Production Facility (Berrow et al., 2007)) using the ligase independent In-Fusion technology. In particular, we selected the dual vectors pOPIN-F and pOPIN-M, which allow the expression both in bacteria and in mammalian cell cultures. pOPIN-F introduces a His_6-tag at the N-terminus of the protein, whereas pOPIN-M adds an N-terminal His_6-tagged maltose binding protein (MBP). Both gene constructs include a PreScission protease site that allows the selective cleavage of the genetic tag. PreScission is a versatile engineered human rhinovirus 3C protease fused to the glutathione S-transferase protein (GST) that works very effectively at low temperature and under conditions of high ionic strength or high glycerol. The usage of this protease is also convenient since it can be produced inexpensively in the laboratory (BOX 3).

```
Human     VAEPELMGTPDGTCYPPPPVPRQASPQNLGTPGLLHPQTSPLLHSLVGQHILSVQQFTKD 1939
E. coli   ----------------------------------------MANPLYQKHIISINDLSRD 19
                                                   : . *  :**:*:::::*

          QMSHLFNVAHTLRMMVQKERSLDILKGKVMASMFYEVSTRTSSSFAAAMARLGGAVLSFS 1999
          DLNLVLATAAKL----KANPQPELLKHKVIASCFFEASTRTRLSFETSMHRLGASVVGFS 73
          ::. :: .* .*    : : . ::** **:** *:*.****  ** ::* ***.:*:.**

          EAT--SSVQKGESLADSVQTMSCYADVVVLRHPQPGAVELAAKH-CRRPVINAGDGVGEH 2056
          DSANTSLGKKGETLADTISVISTYVDAIVMRHPQEGAARLATEFSGNVPVLNAGDGSNQH 133
          :::  *  :***:***::..:* *.*.:*:**** **..**::.  . **:*****  :*

          PTQALLDIFTIREELGTVNGMTITMVGDLKHGRTVHSLACLLTQYRVS-LRYVAPPSLRM 2115
          PTQTLLDLFTIQETQGRLDNLHVAMVGDLKYGRTVHSLTQALAKFDGNRFYFIAPDALAM 193
          ***:***:***:*  * :: : ::******:*******:  *:::  . : ::** :* *

          PPTVRAFVASRGTKQEEFESIEEALPDTDVLYMTRIQKERFGSTQEYEACFGQFILTPHI 2175
          PQYILDMLDEKGIAWSLHSSIEEVMAEVDILYMTRVQKERLDP-SEYANVKAQFVLRASD 252
          *  :  :: .:*   . ..****.: :.*:*****:****:   .**    .**:*

          MTRAKKKMVVMHPMPRVNEISVEVDSDPRAAYFRQAENGMYIRMALLATVLGRF---    2225
          LHNAKANMKVLHPLPRVDEIATDVDKTPHAWYFQQAGNGIFARQALLALVLNRDLVL    311
          : .** :* *:**:***:**:.:**. *:* **:** **:: * **** ** *
```

Figure 8. Sequence alignment of C-terminal human CAD and *E. coli* catalytic subunit. The analysis was performed in the Clustal Omega webserver. The sequences share 44% identity.

BOX 3. Purification of PreScission protease. PreScission protease was induced by the addition of the IPTG at 20 ºC for 16 h. We noticed that to express the protein at low temperature is important, since at 37 ºC the production yield is better but the protein becomes inactive. The band of the overexpressed protein was clearly observed in the supernatant (Figure 9). The protease was purified by two-step produce involving a glutathione column (GSTrap) and an ion-exchange column. To minimize protein degradation, all the purification steps were performed at 4 ºC. The protein eluted from the GSTrap column is around 15 - 20 mg of total protein per liter of cell induced. Although the sample seems to elute almost pure, a second chromatography step was needed. Thus, the sample eluted from the GSTrap was dialyzed for 16 h to change the buffer and increasing the NaCl concentration. PreScission protease did not bind the ion-exchange column and eluted in the flowthrough. The contaminants were retained in the column. Pure protein was concentrated up to 1 mg ml^{-1} and supplement with glycerol prior freezing in liquid nitrogen and stored in 250 µl aliquots at -80 ºC.

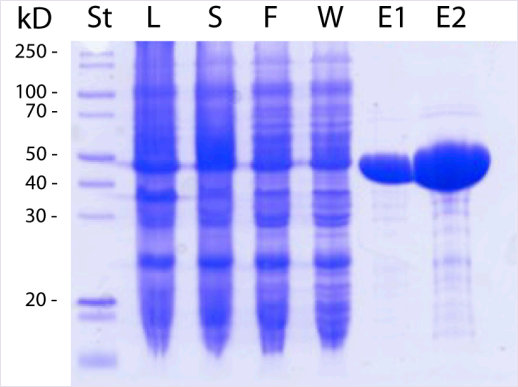

Figure 9. Purification steps monitored by Coomassie-stained SDS-PAGE (12%). St, standards of molecular weight; L, lysate; S, supernatant; F, flowthrough; W, wash; E1, elution from GStrap; E2, elution from ion-exchange.

1.2. Protein expression

We tested the expression of the huATCase constructs in two different *E. coli* strains: Rosetta(DE3) and BL21(DE3). The transformation of pOPIN-F and pOPIN-M vectors, even without any inserted gene, strongly inhibited the growth of lysogenic DE3 bacteria. This effect was palliated by using pLysS strains and media supplemented with glucose, suggesting that the toxicity problem could be related to an undesired strong activation of the lac promoter and leaky protein expression. Both Rossetta pLysS and BL21 pLysS cells produced soluble protein at similar levels and were used at convenience during the course of this work. Alternatively, we also express of the two protein constructs in HEK293 GnTI^{-}cells. This particular cell line has a mutation in the N-acetyl-glucosaminyltransferase I that reduces the branching of polysaccharides in the cell surface and facilitates the growth in suspension (Aricescu et al., 2006). By using either bacteria or HEK293 cells, we observed that the N-terminal fusion to the MBP is key for the expression and solubility of the protein (Figure 10). Since *E. coli* produces larger amounts of recombinant protein than HEK293 cells and it is also faster, cheaper and easier to scale up, we selected bacteria as the system of choice for

the production and isolation of huATCase. Initially, the protein was expressed in bacteria inducing with IPTG (data not shown). Later on, we changed to an autoinduction expression protocol (Studier, 2005) to increase the volume and density of the cultures and thus, obtain larger amounts of recombinant protein.

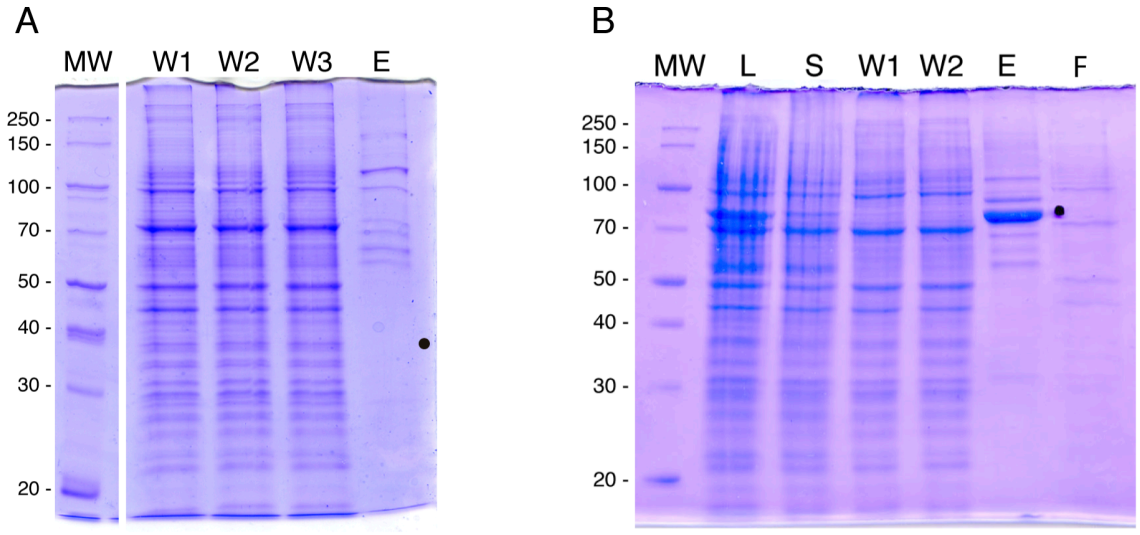

Figure 10. Expression of huATCase in HEK293 cells. Analysis by 12% SDS-PAGE of a small-scale (50ml) production of His_6-huATCase (~38 kDa) **(A)** and His_6-MBP-huATCase (~80 kDa) **(B)**. The cells expressing the protein were lysated (L) and the clarified supernatant (S) was incubated with Ni^{2+}-chelating resin. The non-bound sample (flowthrough, F) and the washes with 10, 30 and 50 mM imidazole (W1–W3) do not contain the protein of interest. The elution (E) of the protein from the beads with 250 mM imidazole shows that only the MBP fused protein is produced. Protein standards are indicated (in kDa). The black dots indicate the expected migration of the huATCase proteins.

1.3. huATCase purification

We devised and optimized a purification protocol to isolate huATCase that included four chromatography steps: Ni^{2+}-affinity, anion-exchange, heparin and size-exclusion (Figure 10A-D). The overexpressed protein was clearly observed by SDS-PAGE both in the lysate supernatant and in the eluted fractions from the Ni^{2+}-affinity column (Figure 11E). The estimated molecular mass of the protein band was around 80 KDa (Figure 10E, lane E_{Ni}), in good agreement with the expected theoretical size (79.85 KDa). At this step, the average yield was of approximately 50 mg of total protein per litre of bacterial culture, out of which we estimated that more than 80% corresponded to MBP-huATCase. The sample was then dialysed overnight to eliminate the imidazole, and at the same time, PreScission protease was added inside the dialysis bag to cleave the His_6-MBP-tag. After removal of the tag, three extra residues remained at the N-terminus of huATCase ($GPMS^{1915}$). The cleaved MBP, which proved to be key for the expression of huATCase, turns into a major contaminant of the preparation (Figure 11E, lane D). The second Ni^{2+}-column step or the usage of an amylose resin were insufficient to get rid of the MBP tag. In turn, the introduction of an anion-exchange chromatography proved to be a very effective purification step. huATCase binds to the column and elutes in a salt gradient nearly pure, with the sole exception of a contaminant of ~20 kDa (Figure 11E, lane E_S). This contaminant is highly soluble compared to huATCase

and due to the losses of huATCase protein during the subsequent concentration steps, the contaminant can reach up to 20 % of the total protein in the sample, which was clearly undesirable for crystallization studies. Mass-spectrometry analysis (carried out at the CNIO Proteomic Unit) identified the contaminant as *E. coli* catabolite activator protein (CAP; also known as cAMP receptor protein; UniProt P0ACJ8), a well-known DNA-binding protein. Thus, we introduced as an additional purification step a heparin column to which CAP bound tighter than huATCase (Figure 11C). After this step, huATCase appeared >98% pure (Figure 11E, lane E_H) and was concentrated and injected onto a size-exclusion chromatography column, the final step of purification. huATCase eluted as a single and sharp peak with no aggregates appearing at the void volume of the column (Figure 11D). The fractions corresponding to the huATCase peak were pooled and concentrated to approximately 3–5 mg ml^{-1}. The final yield of the purification was 10 mg of pure protein per litre of cell culture.

The purified huATCase was initially flash-frozen in liquid nitrogen with no cryo-protectant or supplemented with 40% glycerol and then, stored at -80 ºC. However, we noticed that the specific activity of the protein after three-weeks storage decay up to 2-fold (data not shown). Previous work on purified hamster CAD, reported that the sample was frozen with 30% DMSO (v/v) and 5% glycerol (w/v) (Coleman et al., 1977). This freezing condition proved to be effective also for huATCase, which preserved the activity up to two months after storage at -80 ºC. All the experiments described in this work were performed with freshly purified sample or with protein frozen in DMSO/glycerol and stored at -80 ºC for less than one month.

2. huATCase forms homotrimers in solution

The purified huATCase migrated on a SDS-PAGE at the expected molecular weight of 35 kDa (Figure 11E-F). Contrary, the protein eluted from a size-exclusion column as a single peak at ~14.9 mL (Figure 11E), which according to the calibration with standards of known molecular weight, corresponded to a homotrimer with approximate mass of 110–125 kDa (Figure 12A). To accurately measure the molecular mass of the oligomer, huATCase was injected at 1 mg ml^{-1} onto a size-exclusion column coupled to a multi-angle light scattering device (SEC-MALS). Light scattering measurements along the elution peak indicated the formation of a single species of average molecular mass 99.6 kDa (± 0.2%), in good agreement with the theoretical mass of the expected trimer (104.6 kDa) (Figure 12B).

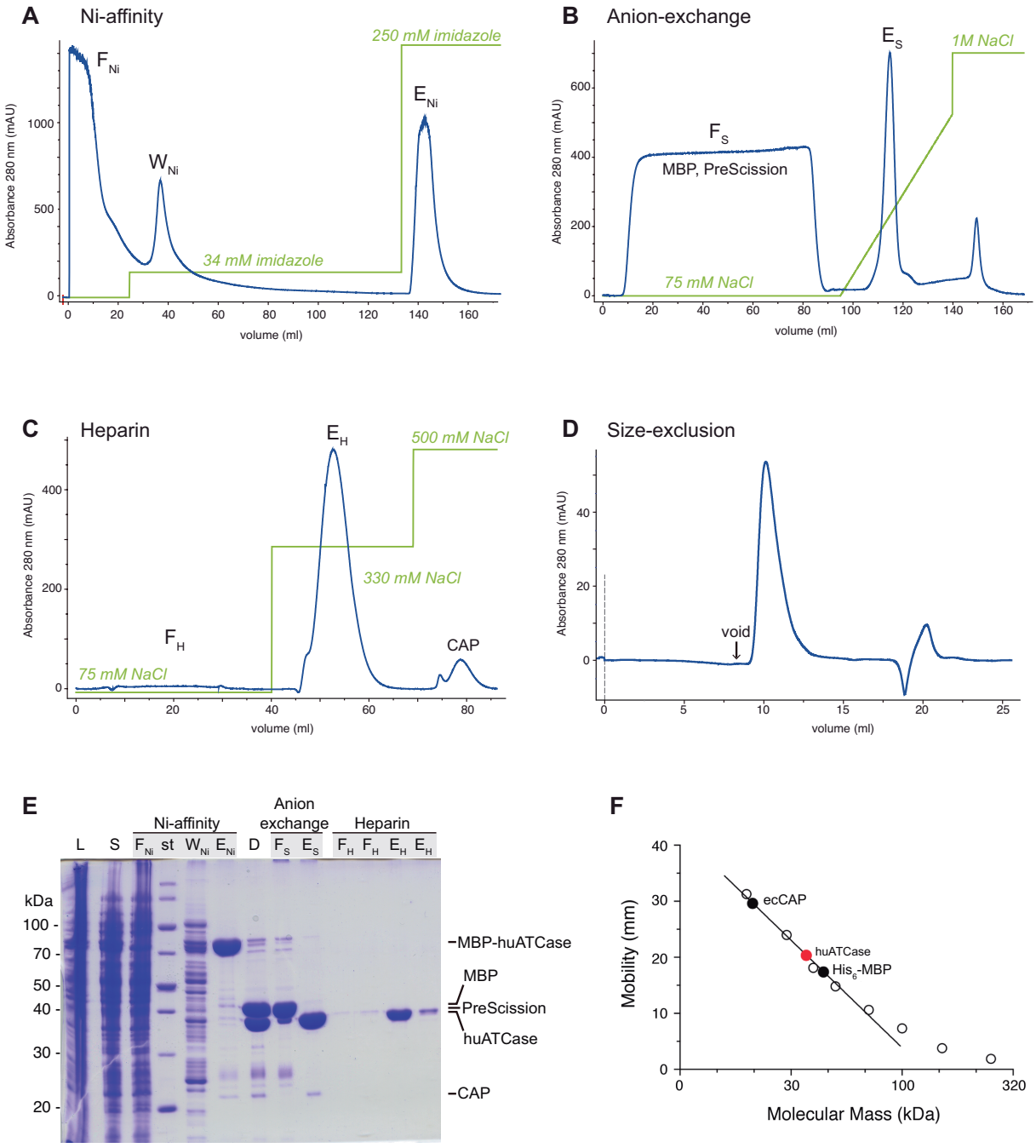

Figure 11. huATCase purification. (A) Chromatogram of Ni^{2+}-chelating HisTrap FF column. The blue line shows the elution profile monitored by UV absorption at 280 nm. The increase of imidazole in the buffer is indicated with a green line. Most contaminants do not bind to the column and flow through (F_{Ni}) or are washed away (W_{Ni}) with 34 mM imidazole. The protein, eluted in 250 mM imidazole (E_{Ni}), was dialyzed overnight in the presence of PreScission to reduce salt concentration and cleave the His_6-MBP tag. **(B)** Chromatogram of a HiTrap SP column. The cleaved His_6-MBP appears in the flowthrough (F_S), whereas huATCase is eluted in a salt gradient (green line) at ~120 mM of NaCl (E_S). **(C)** The eluted protein from HiTrap SP was directly loaded onto a heparin column equilibrated in buffer with 75 mM NaCl. huATCase and CAP bind to the column with different affinities and can be separated using a salt gradient, with huATCase eluting at 330 mM NaCl (E_{Hep}). **(D)** Chromatogram of size-exclusion chromatography (Superdex 75 16/600). The void volume at ~8 ml is indicated. **(E)** Purification steps monitored by Coomassie-stained SDS-PAGE (12%). L, cell lysate; S, clarified supernatant; D, dialyzed and digested with PreScission; other lanes are labelled as in previous panels. Protein molecular-weight markers are indicated on the left of the gel. **(F)** Plot of the electrophoretic mobility of protein markers used to estimate the molecular weight of huATCase. Abscissa axis is represented in log scale. Open circles correspond to molecular-weight markers, closed circles correspond to cleaved His_6-MBP tag and *E. coli* CAP; huATCase is indicated with a red circle.

Interestingly, we noticed that when the protein was injected at concentrations below 0.5 mg ml^{-1}, the elution peak showed a shoulder at ~16.9 ml. Initially, this was believed to correspond to the contaminating CAP protein, but as the purification protocol was optimized, the hypothesis that it might correspond to partial dissociation of the trimer gained weight. To test this possibility, the fractions from the first half of the peak corresponding to the trimer were pooled and re-injected in the size-exclusion column, and this procedure was repeated a second time. We observed that as the concentration of the re-injected sample decreased, the shoulder peak became more apparent (Figure 12C-D). The MALS analysis of this shoulder peak indicated a molecular mass of 39.7 (± 0.64%), in agreement with the expected size of the huATCase monomer.

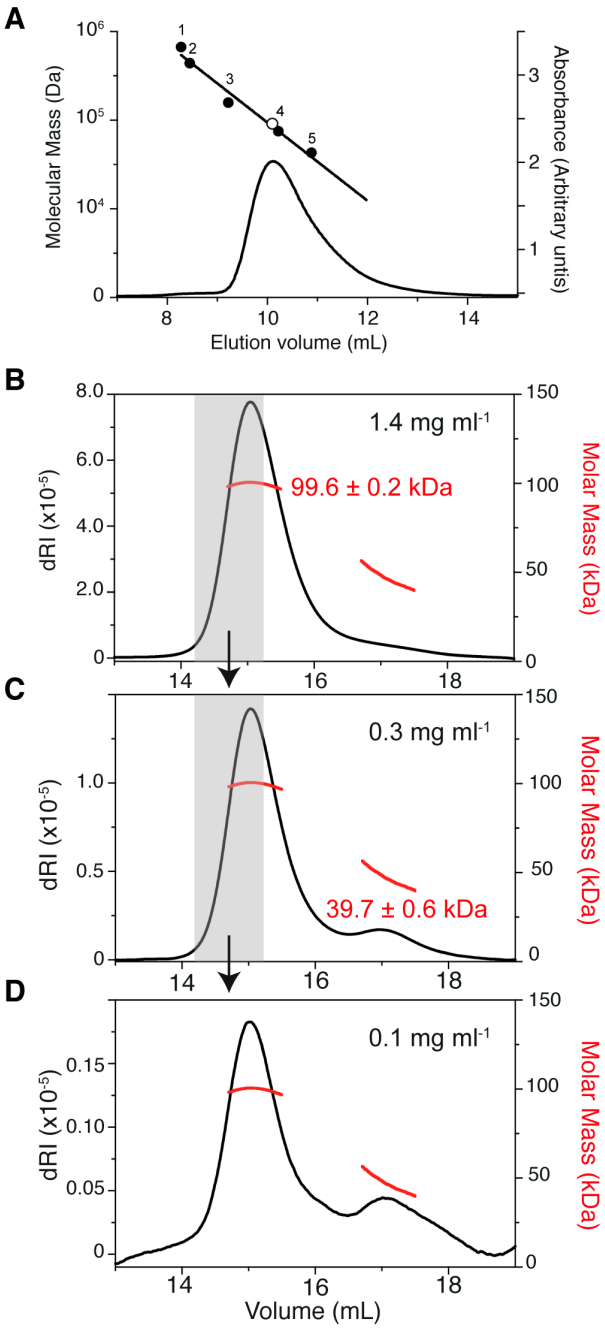

Figure 12. Determination of the oligomeric state of huATCase by SEC-MALS. (A) Semilogaritmic plot of the molecular mass versus the elution volume from a Superdex 75 gel filtration column. Closed circles indicate the elution volume of proteins standards (1, thyroglobulin; 2, ferritin; 3, aldolase; 4, conalbumin; 5, ovalbumin). The empty circle indicates the expected position for a huATCase homo-trimer (theoretical mass 105 kDa). The black trace shows the elution profile of huATCase from the column. **(B)** SEC-MALS analysis using a Superdex 200 column. Protein elution profile with the measurement of the differential refractive index (dRI; black trace; left axis) and the corresponding molar mass (red dots; right axis). The grey shades indicate the fractions of the peak that were pooled and re-injected in the column **(B,C)**. The concentrations at which the protein was injected in the columns were: 1.4 mg ml^{-1} **(B)**, 0.3 mg ml^{-1} **(C)**, and 0.1 mg ml^{-1} **(D)**.

We confirmed the dissociation of the huATCase trimer by analytical ultracentrifugation (AUC). Sedimentation velocity assays demonstrated the presence of two species with different hydrodynamic size and sedimentation coefficients (Figure 13A-B). The most prominent species, sedimenting at 5.2 S, corresponded to a trimer of ~105 kDa, whereas a second species of 2.7-3 S agreed with a protein monomer of ~35 kDa. We measured the relative proportion of the monomer versus

trimer at different protein concentrations. The most diluted sample that could be accurately measured with the instrument had a concentration of 0.03 mg ml^{-1} (0.86 µM) and exhibited 79 % of monomer. Plots of the relative percentage of trimer versus protein concentration fitted to a hyperbola from which we estimated a dissociation constant (K_D) of 0.27 ± 0.04 µM (Figure 13C).

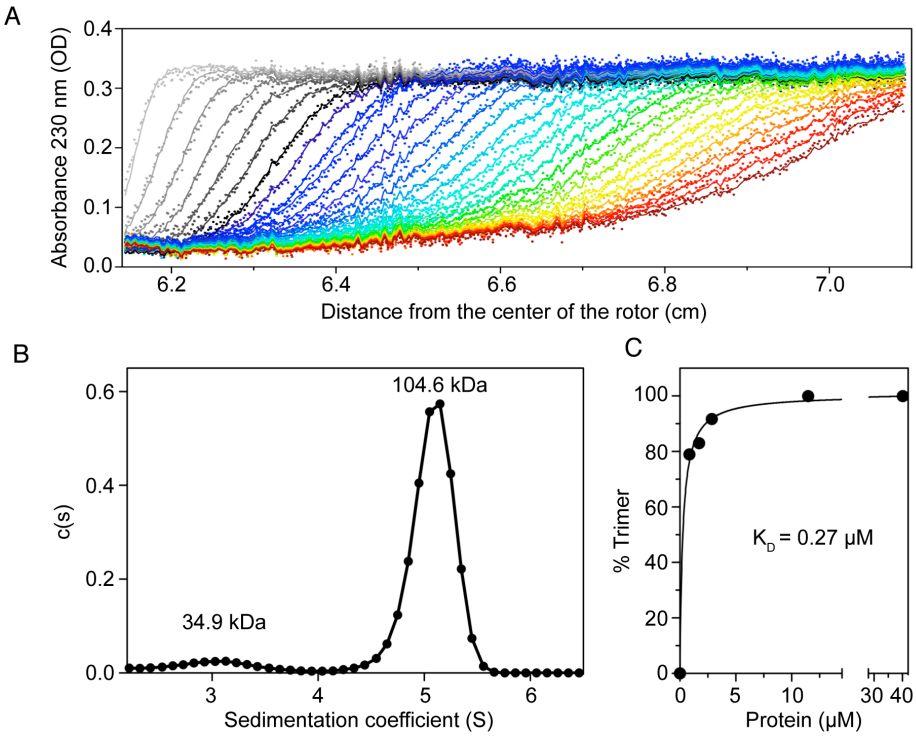

Figure 13. Measurement of huATCase dissociation constant. (A) Sedimentation velocity profile at 0.1 mg ml^{-1} huATCase measured by absorbance at 232 nm. The absorbance scan across the centrifuge cell is shown in different colors. In the first scan (grey color), the sedimentation of huATCase has already depleted its concentration in the region near the top of the tube and formed a sedimentation boundary. As the experiment progresses, the depleted region expands and the boundary moves away from the center of the rotor. By the end of the experiment (red color), the concentration of protein has dropped to essentially zero throughout the upper half of the cell. **(B)** Transformation of the data in **(A)** into a sedimentation coefficient distribution. The graph resembles a size-exclusion chromatogram (except that the peaks come in the opposite order), where the area under the peaks gives the relative amount of each species present in the sample. The apparent diffusion coefficient, calculated from the peak width, and the ratio of sedimentation coefficient to diffusion coefficient allow estimating the molecular mass of each species. **(C)** Graphical representation of the percentage of trimer measured at different protein concentrations. The data fit to a hyperbola (black line) from which a dissociation constant (K_D) can be estimated.

3. huATCase stability is enhanced by the binding of ligands

We used differential scanning fluorimetry (thermofluor) to test the stability of the protein at different pH values and with different ligands (Niesen et al., 2007). The temperature at which huATCase unfolds was measured in a quantitative PCR machine by the increase in fluorescence of SYBR orange that binds to exposed hydrophobic areas of the protein. In the absence of ligands, the transition temperature or temperature of melting (T_M) of huATCase was within 50.6 – 54.8 ºC range at a pH between 6 and 9 (Figure 14A). Interestingly, the stability of the protein increased notably in phosphate buffer (T_M = 70 ºC),

likely due to the binding of inorganic phosphate to the active site. Indeed, we observed that the stability of huATCase increased when CP was added to the sample, up to maximum T_M of 70 ºC (Figure 14B). Titration of PALA caused an even higher stabilization of the huATCase, with a rise of the T_M up to 83 ºC. Considering this increase in stability as a rough indicator of the binding of the ligands to the protein we estimated a high affinity for CP and PALA, with dissociation constants (K_D) below 1 µM (Figure 14C).

Importantly, the addition of Asp in the absence of CP did not affect the stability of the protein (Figure 14B), suggesting that, as reported for other ATCases (Porter et al., 1969), the human enzyme also follows a compulsory order in the binding of the substrates. The addition of Asp to huATCase already bound to CP cannot be measured since the reaction would take place. Thus, we titrated an inert Asp analogue, succinate (Porter et al., 1969), to a pre-incubated mixture of huATCase and CP. The combination of CP and succinate did not mimic the stabilizing effect of PALA (Figure 14B), indicating that the mixture of both ligands does not trigger a conformational change comparable to that with PALA, which likely explains the high stabilization of the protein.

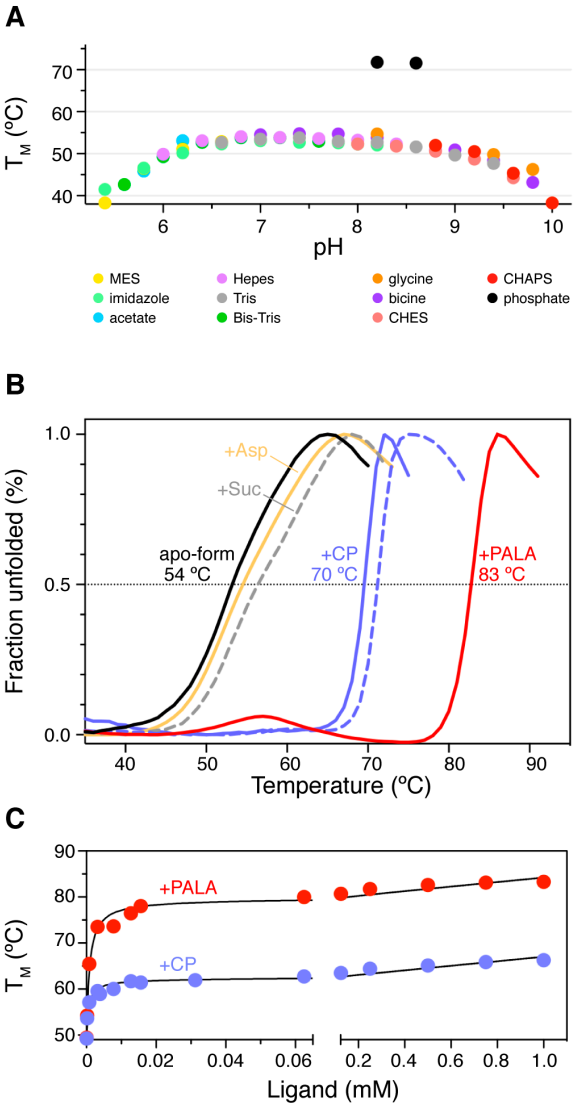

Figure 14. huATCase stability and ligand binding assessed by differential scanning fluorimetry. (A) huATCase melting temperature (T_M) monitored by thermal shift assays at different buffer conditions. **(B)** Denaturing curves measured in buffer Tris-acetate pH 8.3 in the absence and presence of CP, Asp, succinate (Suc) and PALA. **(C)** Graphical representation of the increase in T_M at varying concentration of CP or PALA.

4. huATCase exhibits positive cooperativity for the binding of substrates

The enzymatic activity of the isolated huATCase domain was determined by measuring the production of carbamoyl aspartate (CAsp) in a discontinuous colorimetric assay (Prescott and Jones, 1969) adapted to a 96-well plate format (Else and Hervé, 1990). Due to the low stability of CP when it is not bound to the enzyme ($t_{1/2}$ of ~1h at 25 ºC) (Legrain et al., 1995), the reaction was started by the addition of CP to a pre-incubated mixture of huATCase and Asp (Stebbins et al., 1989). First, we estimated the initial rate and ensured that it was proportional to the total enzyme concentration. For this, we performed time-course experiments at 37 ºC with three different concentrations of huATCase and 10 mM Asp and 5 mM CP (Figure 15A). The production of CAsp, calculated by interpolation in a standard curve prepared with L-CAsp (Figure 15B), progressed linearly up to approximately the consumption of 30 % of the CP in the reaction. According to these data, the reaction rate was proportional to the concentration of enzyme used and the velocity was calculated as 8 mmol of CAsp produced per h and per mg of protein (Figure 15C). Further enzymatic assays were carried out with 0.04 µM of enzyme and stopped after 5 min. In addition, the temperature of the experiment was lowered to 25 ºC to compare the huATCase activity with those of bacterial homologues, mostly measured at this temperature (Stebbins et al., 1989).

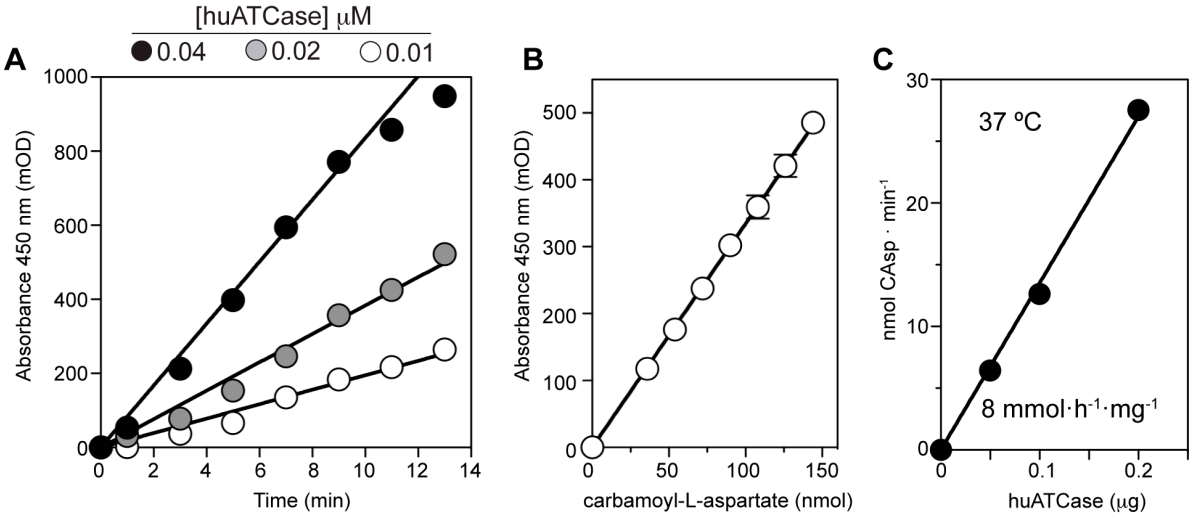

Figure 15. huATCase activity. (A) Time-course experiment at three different huATCase concentrations. **(B)** Standard curve of the absorbance of known concentrations of carbamoyl-L-aspartate (CAsp). **(C)** The production of CAsp increases linearly with the concentration of enzyme.

Plots of initial rates at 25 ºC, measured at fix concentration of Asp (10 mM) and varying concentration of CP, followed a sigmoidal curve that fitted to the Hill equation with V^{app}= 4.1 mmoles h^{-1} mg^{-1}, $[CP]_{0.5}^{app}$ = 0.2 mM and n_H=1.6 (Figure 15A). The deviation from the Michaelis-Menten equation was better diagnosed in an Eadie-Hofstee plot (Figure 16A, inset). On the other hand, initial rates at 5 mM CP and varying concentration of Asp followed a complex trend (Figure 16B). Up to 7.5 mM Asp, the data fit a sigmoid with V^{app} = 4.5

mmoles h^{-1} mg^{-1}, $[Asp]_{0.5}^{app}$ = 2.2 mM and n_H = 1.3, whereas at higher concentrations, Asp has a pronounced inhibitory effect (Figure 16B, inset).

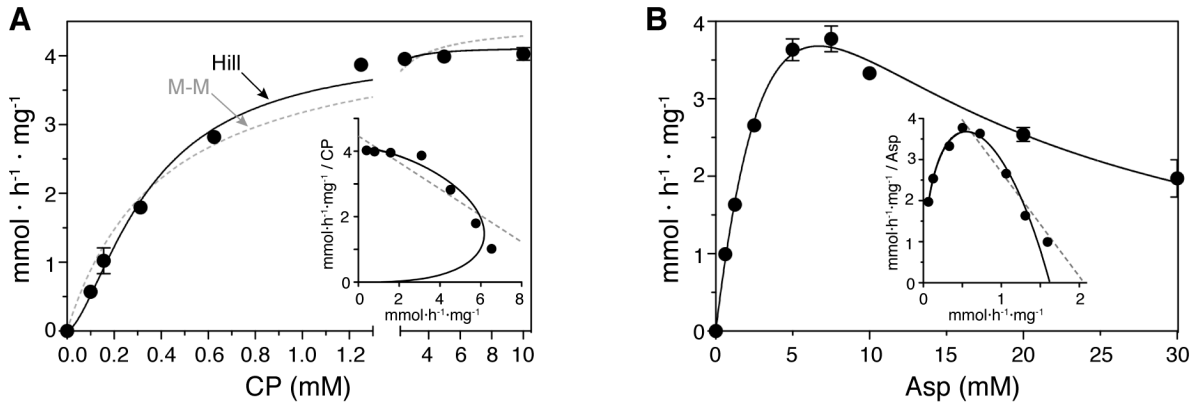

Figure 16. Substrate saturation curves. (A) huATCase activity measured at 25 ºC by fixing the concentration of Asp (10 mM) and varying the concentration of CP. Fitting to a Hill equation (sigmoid; solid line) is more accurate than to a Michaelis-Menten equation (hyperbola; dashed grey line). **(B)** Activity assays at fixed concentrations of CP (5 mM) and varying the concentration of Asp. The data was fitted taking into account the inhibition at high Asp concentrations. The insets in both graphs show the Eadie-Hoffstee plot of the data to highlight the deviation from the Michaelis-Menten behaviour (dashed grey line).

Kinetic values of V^{app} and $S_{0.5}$ for huATCase are in the same order of magnitude to those observed for the isolated ATCase domain from hamster CAD (Qiu and Davidson, 1998, Grayson and Evans, 1983). However, sigmoidicity in the kinetic curves indicates that huATCase exhibits positive cooperativity for the substrates. This means that binding of CP and Asp at one active site increases the substrate affinity and/or activity in the other subunits within the trimer. The cooperativity effect in huATCase was unexpected since the isolated catalytic trimer of bacterial ATCases lack the homotropic properties and are reported to be non-cooperative (Gerhart and Holoubek, 1967, Porter et al., 1969, Stebbins et al., 1989).

To further test the cooperativity effects on huATCase, we performed activity assays at sub-saturating Asp concentration in the presence of increasing concentrations of the competitive inhibitor succinate. It has been shown in in *E. coli* ATCase (Stebbins et al., 1989) that low succinate concentrations stimulate the activity by promoting the T-R transition, whereas at higher concentrations, succinate has an inhibitory effect. Contrary to ecATCase, we did not detect stimulation of the activity in presence of succinate (Figure 17). Succinate inhibits huATCase activity at the three Asp concentrations that were tested.

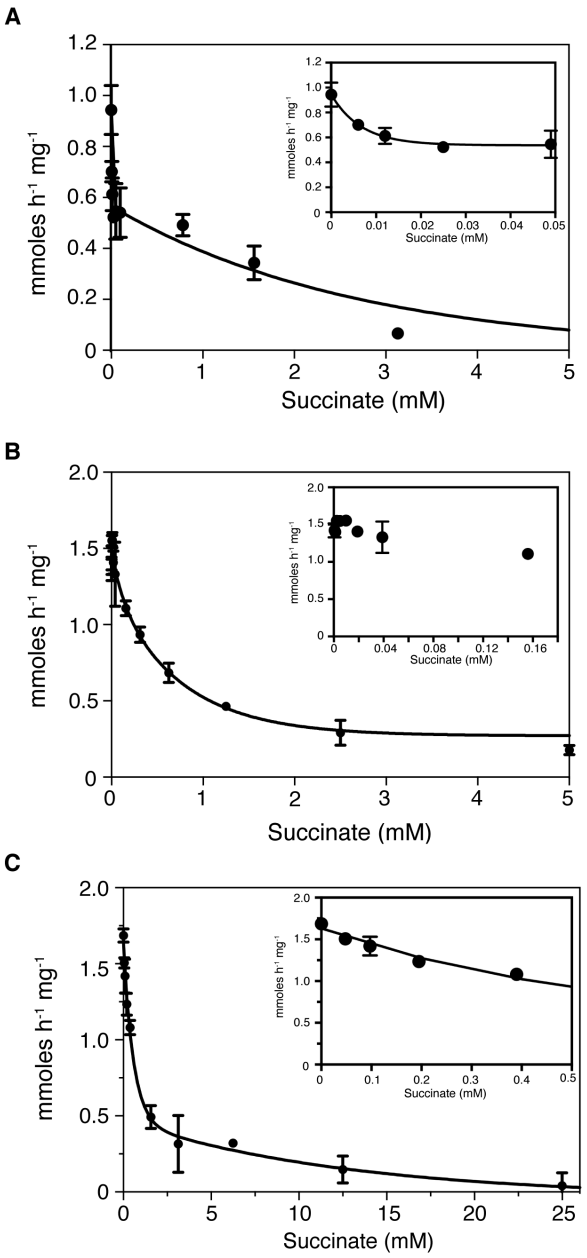

Figure 17. huATCase activity in the presence of succinate. Activity measured at 25 ºC at increasing concentrations of succinate in 5 mM CP and the following sub-saturating concentrations of Asp: **(A)** 0.2 mM, **(B)** 0.5 mM and **(C)** 1 mM. Insets, details at low succinate concentrations.

5. huATCase binds PALA with negative cooperativity

5.1. PALA strongly inhibits huATCase activity

To test the effect of PALA on huATCase activity, we performed substrate saturation curves triggering the reaction with CP at increasing concentrations of the inhibitor (Figure 18A,B). huATCase activity is inhibited by PALA at substoichiometric concentrations -lower concentrations than those of the enzyme used in the reaction–, which agrees with the tight binding of the inhibitor described for *E. coli* ATCase (Collins and Stark, 1971, West et al., 2008). The inhibition by PALA could not be overcome by raising the concentrations of any of the two substrates (Figure 18A,B), as if PALA was a non-competitive inhibitor of the reaction.

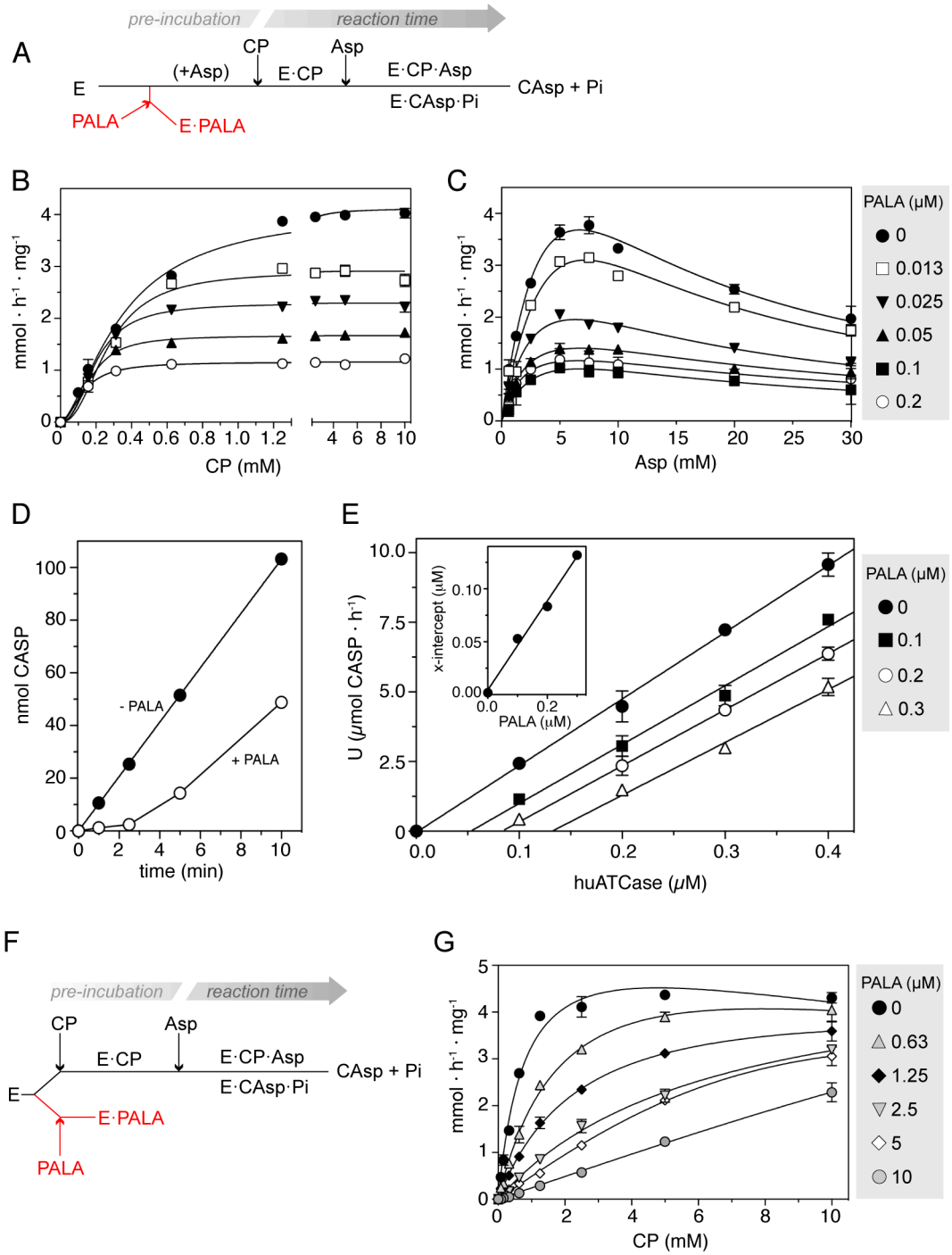

Figure 18. huATCase inhibition by PALA. (A) Cleland-diagram of the experimental setup. huATCase is pre-incubated with Asp and PALA and the reaction is started by addition of CP. **(B,C)** Saturation curves at increasing concentrations of PALA. **(D)** Time-course experiments in absence and presence of PALA. **(E)** Ackerman-Potter plot representing the activity of huATCase mixed with approximately stoichiometric concentrations of PALA. The prolonged linear section of the curves intersects with the abscissa axis, indicating the concentration of enzyme titrated out by PALA during the reaction, as represented in the inset. **(F)** Cleland diagram showing a different experimental setup, where the reaction is triggered by the addition of Asp. **(G)** CP saturation curves at increasing concentrations of PALA following the experimental setup in **(F)**.

Time-course activity experiments in presence of PALA showed a lag in the linear formation of CAsp (Figure 18D). This indicated that, although the binding of PALA is reversible, its dissociation is a slow process compared to the binding of the substrates or the reaction turnover. Thus, PALA behaves as a “slow-tight binding inhibitor” (Morrison, 1969,

Morrison, 1982). Rather than being a non-competitive inhibitor, PALA binding sequesters the enzyme in an “inactive state” during most of the reaction time, causing an apparent reduction in V^{app}, whereas the kinetic properties ($S_{0.5}^{app}$, k_{cat}) of the free enzyme remained unchanged.

To further test whether PALA is a slow-tight binding inhibitor, we incubated huATCase and PALA at similar concentrations and measured the residual activity. The Ackermann-Potter plot in Figure 17E confirmed that PALA is a substoichiometric inhibitor, and the intersections of the prolonged linear portions of the curves with the abscissa axis indicate the amount of enzyme titrated out by the inhibitor during the reaction (Ackermann and Potter, 1949).

As a rough approximation to the steady-state, we pre-equilibrated the enzyme with CP and PALA and then, started the reaction by adding Asp (Figure 18F). Using this experimental set up, we observed that inhibition occurred at concentrations of PALA ~100-fold higher than in the previous experiments. In this case, the inhibition was overcome by the increasing concentrations of CP, demonstrating that indeed, CP and PALA compete for the binding to the same form of the enzyme. Analysis of the data yielded a K_i^{PALA} = 351 nM. This high value is at odds with the tight stoichiometric binding, suggesting that this approach is not valid to quantify the interaction between huATCase and PALA.

Another intriguing result was that we failed to completely inhibit huATCase activity at the PALA concentrations tested (Figure 18B,C,G). This was likely due to the slow interchange of PALA with the large excess of CP present in the solution during the reaction time (Figure 18B). To attain full inhibition, we increased the concentration of PALA up to 100 µM (Figure 19A). Notably, huATCase activity decreased with PALA concentration following a triple exponential decay, from which we estimated I_{50} values of 0.014 µM, 1.4 µM and 14 µM. This result was again in conflict with the knowledge about bacterial ATCases catalytic trimers, since it strongly suggested that huATCase trimer exhibits different affinities for the binding of the inhibitor.

5.2. Quantification of CP and PALA binding by isothermal titration calorimetry

To accurately measure the binding of substrates and PALA to huATCase, we performed isothermal titration calorimetry (ITC) assays. We proved the existence of three independent and equivalent binding sites for CP per huATCase trimer, with a K_D^{CP} = 6.3 µM (Figure 19B). We failed to detect the binding of Asp or succinate to the protein alone. We neither observed the binding of succinate in a pre-incubated mixture of enzyme succinate. When higher concentrations of succinate were tested, the heat released by small differences in the pH between the sample in the cell and the injected solution of ligand masked any possible changes in enthalpy caused by the binding of the ligand (data not shown).

Surprisingly, the titrations with PALA proved the existence of three non-equivalent binding sites per huATCase trimer with increasing intrinsic dissociation constants K_D^1 = 0.017

μM, K_D^2 = 0.027 μM and K_D^3 = 1.4 μM (Figure 19C). This negative cooperativity for the binding of PALA, with nearly 100-fold difference in affinity between the first and third binding sites, was unexpected, since, as already mentioned, the isolated catalytic trimers of bacterial ATCases lack homotropic effects between the active sites. Further clarification required a detailed characterization of the huATCase structure.

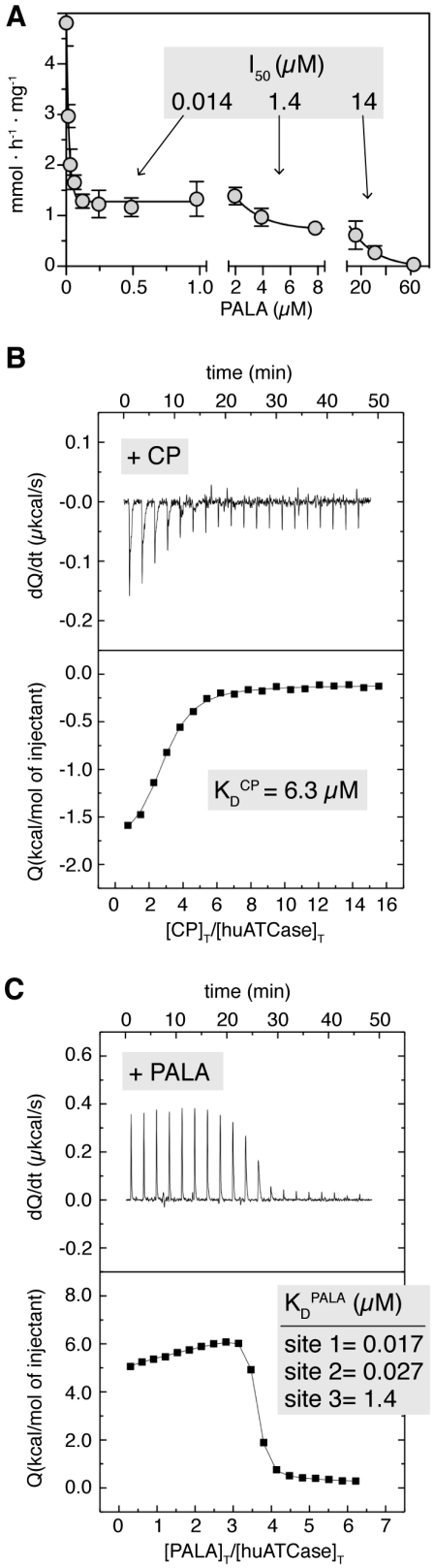

Figure 19 Measurements of PALA binding to huATCase. (A) Inactivation of huATCase at increasing concentrations of PALA follows a triple exponential decay. I_{50} values are indicated. **(B,C)** Calorimetric titrations for the interaction of huATCase with CP **(B)** or PALA **(C)**. Upper plots show the thermogram (thermal power as a function of time) and lower plots show the binding isotherm (normalized heat per injection as a function of the molar ratio ligand:protein). Non-linear fitting analysis was performed considering three identical non-cooperative binding sites **(B)** or three non-identical cooperative binding sites **(C)**. Intrinsic site-specific dissociation constants are indicated.

6. huATCase crystallization

6.1. Initial crystallization screening

huATCase crystallized at 18 ºC under a variety of conditions tested from commercial screenings. Most crystals appeared in conditions with 10-25 % polyethyleneglycol of different sizes as precipitant, 100-200 mM salt (most frequently $MgCl_2$) and at pH values between 5 and 9. The conditions yielded a large number of small crystals that usually appeared overnight. The crystals were thin plates with approximate dimensions of 50 x 50 μm, that most often clustered together. We exhaustively tried to reduce the nucleation and to increase the size and thickness of the crystals. However, neither the changes in the crystallization solutions, temperature or protein concentrations improved the number, size or morphology of the crystals. In turn, we found out that using freshly purified protein significantly reduced the amount of nucleation and resulted in slightly thicker crystals. On the contrary, the use of protein frozen before or after the gel-filtration step yielded poorer results in

the crystallization trials. Therefore, we purified and used it immediately to perform the crystallization experiments.

6.2. Crystallization of huATCase free of ligands

Starting from the condition number 16 of the JCSG+ Suite as the initial hit, we optimized the crystallization of huATCase in absence of ligands to a final condition that contained 0.1 M Tris- HCl pH 8.5, 6% PEG 8000, 9% ethylene glycol (EG). The crystals were optimized in sitting drops by mixing equal volumes of reservoir and of freshly purified protein at 1.5 mg ml^{-1}. The crystals appeared after 1–2 days at 18 ºC and grew without clustering (Figure 20A). Most of the crystals were thin plates with trapezoidal shape and 0.2 mm in the maximum dimension. In the same drops, similar thin but hexagonal-shaped crystals also appeared. These crystals exhibited concave faces, which has been suggested to be a possible indicator of crystal twinning (Yeates, 1997). Prior to cooling in liquid nitrogen, the trapezoidal crystals were transferred to a cryo-protectant solution supplemented with 30% EG.

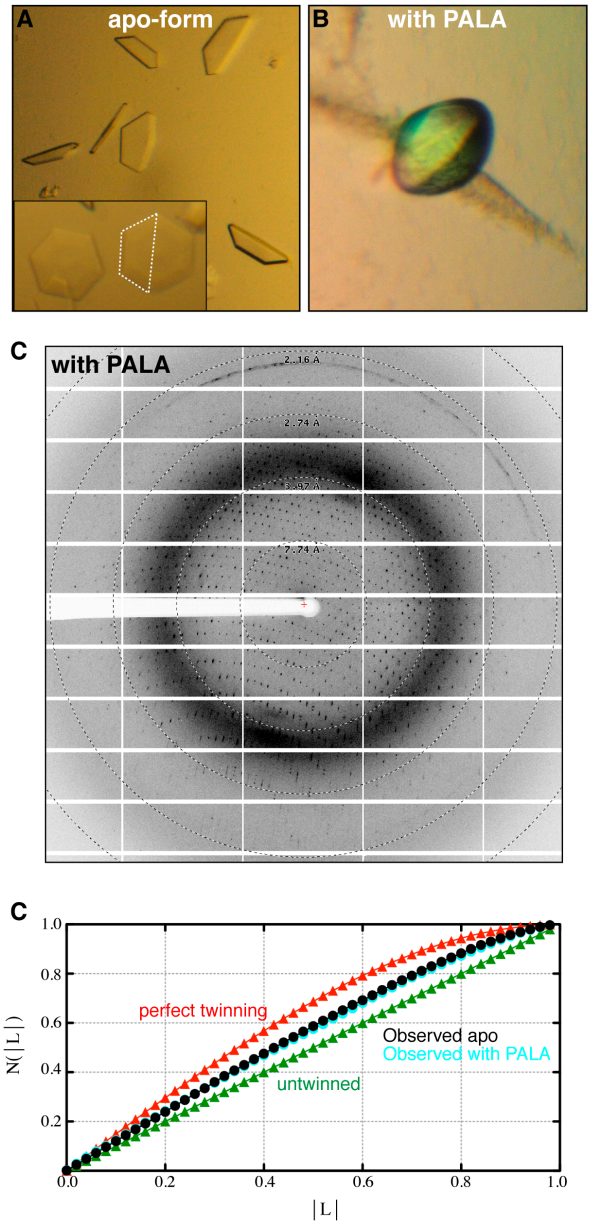

Figure 20. Crystallization of huATCase and X-ray data collection. (A) Crystals of huATCase grown by the hanging-drop method in absence of ligands. The inset shows hexagonal-shaped crystals with an apparent concave face that appeared together with the more common trapezoidal-shaped crystals. The dotted line indicates the relation between the crystals types. **(B)** Rugby-ball-shaped crystals of huATCase grown in the presence of PALA. **(C)** X-ray diffraction pattern of huATCase crystal grown with PALA obtained from the XALOC beamline (ALBA). The circles indicate the resolution limits in Å. **(D)** Twinning analysis of the X-ray diffraction data of the crystals free or grown with PALA processed in space group $P2_1$. The L-test for acentric data indicates a deviation from the expected distribution for a normal untwinned crystal (Padilla and Yeates, 2003).

6.3. Crystal soaking in CP

To obtain the crystal of the complex with CP, we used the method of soaking the apo crystals in mother liquor supplemented with CP. The incubation with the ligand had a visible damaging effect on the crystals that eventually dissolved. We attempted to optimize the concentration of the substrate and the incubation time to preserve the integrity of the crystals. By restricting the incubation time to 5 min in the solution containing 5 mM CP, some larger crystals appeared to be more resistant. These crystals were cryo-protected with solutions of increasing concentration of EG containing CP and flash-frozen for further diffraction studies.

6.4. Co-crystallization of huATCase with PALA

To crystallize the protein in complex with PALA, we first mixed huATCase at 1.5 mg ml^{-1} with 0.5-1 mM PALA in buffer at pH 8. Under these conditions, a precipitated appeared after 30 min, and the remaining soluble protein fraction, at a concentration of 0.5 mg ml^{-1}, did not form crystals. Since the reported crystals of bacterial ATCases in complex with PALA were mostly obtained in conditions with extreme acidic or basic pH, we changed the pH of the protein sample by adding buffer at pH 10 (CAPS) or 5.5 (MES) and then added PALA to a final concentration of 0.5 mM. In this way, the addition of PALA did not cause protein precipitation and the sample could be used for crystallization. Initial crystals of huATCase with PALA were obtained at 18 ºC in condition number 12 of the PACT Suite [0.1 M sodium acetate pH 5, 10 mM zinc chloride, 20% (w/v) PEG 6000]. Optimization of this condition allowed us to grow rugby-ball-shaped crystals with a maximal dimension of 80 µm (Figure 20B). The final crystallization conditions were 50 mM sodium acetate pH 4.8, 10 mM zinc chloride, 15% PEG 6000, mixing huATCase at 3 mg ml^{-1} in gel-filtration buffer with an equal volume of 0.1 M CAPS pH 7, 150 mM sodium chloride and 1 mM PALA. The crystals were cryo-protected with 20% glycerol and flash-frozen in liquid nitrogen.

7. huATCase structural determination

7.1. X-ray diffraction data collection and processing

We used the in-house diffractometer to test that the crystals were indeed made of protein and that diffracted X-rays. All subsequent diffraction experiments were carried out using synchrotron radiation at SLS (Villigen), DESY (Hamburg) and ALBA (Barcelona). In general, the crystals diffracted X-rays poorly. The diffraction spectra were anisotropic and reached a maximum resolution of 4 Å in the best direction. A screening of approximately hundred crystals was needed to collect complete datasets of huATCase free of ligands (huATCase-apo), soaked with CP (huATCase-CP) or co-crystallized with PALA (huATCase-PALA) at resolutions better than 3.2 Å (Figure 20C).

Initial indexed reflections indicated that the three crystal types belonged to the hexagonal space group $P6_322$, with similar unit-cell parameters a = b = 83 Å and c = 145–158 Å. Matthews coefficient calculation predicted the presence of one huATCase protomer per asymmetric unit, with a V_M of 2.35 $Å^3$ Da^{-1} and a solvent content of 48%. Therefore, the threefold axis of the huATCase trimer was coincident with the crystallographic symmetry axis. However, analysis of the diffraction data with phenix.xtriage from the PHENIX package (Adams et al., 2010) indicated a significant deviation of the intensity statistics from the expected distribution and suggested the over-merging of pseudo-symmetric intensities. Reprocessing the diffraction data in lower symmetry space groups yielded significantly better statistics (Tables 4 and 5). It was possible that the crystals belonged to the monoclinic space group $P2_1$, with the peculiarity that the β angle was 120º. The predicted content of the monoclinic asymmetric unit was six huATCase subunits that would form two trimers. This unusual $P2_1$ unit-cell geometry allows pseudomerohedral twinning. Indeed, a subsequent L-test analysis (Padilla and Yeates, 2003) of the X-ray diffraction intensity statistics of the three datasets diagnosed a deviation from the expected distribution of a normal untwinned crystal (Figure 20D).

Table 1. Data-processing statistics of huATCase-apo in different possible space groups.

Crystal system	**Hexagonal**		**Trigonal**			**Orthorombic**	**Monoclinic**	
Unit cell parameters a, b, c (Å) α, β, γ (°)	82.8, 82.8, 158.0 90, 90, 120					82.7, 143.4, 157.9 90, 90, 90	82.7, 143.4, 157.9 90, 90, 90	82.8, 157.9, 82.7 90, 120, 90
Space group	**P622**	**P6**	**P312**	**P321**	**P3**	**C222**	**C2**	**$P2_1$**
Multiplicity	12	6	6	6	3	8	4	2
Reflections (observed/unique)[a] overall outer shell	337,844/19,442 15,320 / 1,357	346,615/35,711 15,282 / 2,576	341,745/36,488 15,325 / 2,606	341,802/37,274 15,327 / 2,606	346,036/70,518 15,261 / 5,023	342,155/ 52,757 15,142 / 2,753	345,170/101,736 15,154 / 5,147	344,600/101,090 15,125 / 5,053
Multiplicity[b]	17.4 (11.3)	9.7 (5.9)	9.4 (5.9)	9.2 (5.9)	4.9 (3.0)	6.5 (5.5)	3.4 (2.9)	3.4 (3.0)
Completeness (%)[b]	99.6 (96.6)	99.7 (97.1)	99.6 (96.8)	99.5 (95.2)	99.5 (95.2)	95.8 (68.9)	95.0 (65.7)	94.3 (64.4)
R_{meas}[b]	**0.179 (0.568)**	**0.123 (0.524)**	**0.188 (0.585)**	**0.189 (0.581)**	**0.120 (0.542)**	**0.191 (0.569)**	**0.204 (0.593)**	**0.115 (0.513)**
I/σ_I[b]	12.5 (4.2)	13.9 (3.8)	8.8 (2.9)	8.7 (2.9)	10.2 (2.7)	7.3 (2.9)	5.0 (2.0)	8.9 (2.8)
$CC_{1/2}$[c]	98.9 (88.5)	99.5 (87.1)	98.7 (80.4)	98.6 (78.2)	99.1 (74.9)	98.1 (76.9)	96.3 (63.4)	98.9 (78.4)

[a]Overall resolution range = 42.4-2.1 Å; outer shell= 2.15-2.10 Å.
[b]Data of the highest resolution shell are shown in parentheses.
[c]Percentage of correlation between intensities from random half-datasets.

Table 1. Data-processing statistics of huATCase-apo in different possible space groups.

Crystal system	**Hexagonal**		**Trigonal**			**Orthorombic**	**Monoclinic**	
Unit cell parameters a, b, c (Å) α, β, γ (°)	82.8, 82.8, 158.0 90, 90, 120					82.7, 143.4, 157.9 90, 90, 90	82.7, 143.4, 157.9 90, 90, 90	82.8, 157.9, 82.7 90, 120, 90
Space group	**P622**	**P6**	**P312**	**P321**	**P3**	**C222**	**C2**	**$P2_1$**
Multiplicity	12	6	6	6	3	8	4	2
Reflections (observed/unique)[a] overall outer shell	337,844/19,442 15,320 / 1,357	346,615/35,711 15,282 / 2,576	341,745/36,488 15,325 / 2,606	341,802/37,274 15,327 / 2,606	346,036/70,518 15,261 / 5,023	342,155/ 52,757 15,142 / 2,753	345,170/101,736 15,154 / 5,147	344,600/101,090 15,125 / 5,053
Multiplicity[b]	17.4 (11.3)	9.7 (5.9)	9.4 (5.9)	9.2 (5.9)	4.9 (3.0)	6.5 (5.5)	3.4 (2.9)	3.4 (3.0)
Completeness (%)[b]	99.6 (96.6)	99.7 (97.1)	99.6 (96.8)	99.5 (95.2)	99.5 (95.2)	95.8 (68.9)	95.0 (65.7)	94.3 (64.4)
R_{meas}[b]	**0.179 (0.568)**	**0.123 (0.524)**	**0.188 (0.585)**	**0.189 (0.581)**	**0.120 (0.542)**	**0.191 (0.569)**	**0.204 (0.593)**	**0.115 (0.513)**
I/σ_I[b]	12.5 (4.2)	13.9 (3.8)	8.8 (2.9)	8.7 (2.9)	10.2 (2.7)	7.3 (2.9)	5.0 (2.0)	8.9 (2.8)
$CC_{1/2}$[c]	98.9 (88.5)	99.5 (87.1)	98.7 (80.4)	98.6 (78.2)	99.1 (74.9)	98.1 (76.9)	96.3 (63.4)	98.9 (78.4)

[a]Overall resolution range = 42.4-2.1 Å; outer shell= 2.15-2.10 Å.
[b]Data of the highest resolution shell are shown in parentheses.
[c]Percentage of correlation between intensities from random half-datasets.

7. 2. Structure determination

Crystallographic phases were obtained by molecular replacement using the catalytic chain of *E. coli* ATCase (PDB 2ATC) as the search model. Initially, we carried out the molecular replacement considering that the crystal belonged to the space group $P6_322$, and we subsequently attempted to build and refine a model for one protein subunit within the asymmetric unit. Although the electron density maps allowed the tracing of most of the protein backbone, the refinement did not progress beyond unusually high values of R factor and R_{free} that did not match the overall decent quality of the electron density maps. By reducing the symmetry of the space group to $P2_1$, we built and refined the structure of two protein trimers per asymmetric unit (Figure 21A). During the initial refinement cycles, non-crystallographic symmetry restraints were maintained between the six subunits. The restraints were gradually released during the refinement as small differences were appreciated between the different subunits (Figure 21B). Even before the refinement was completed, it appeared clear that the huATCase molecule lacked a strict 3-fold symmetry.

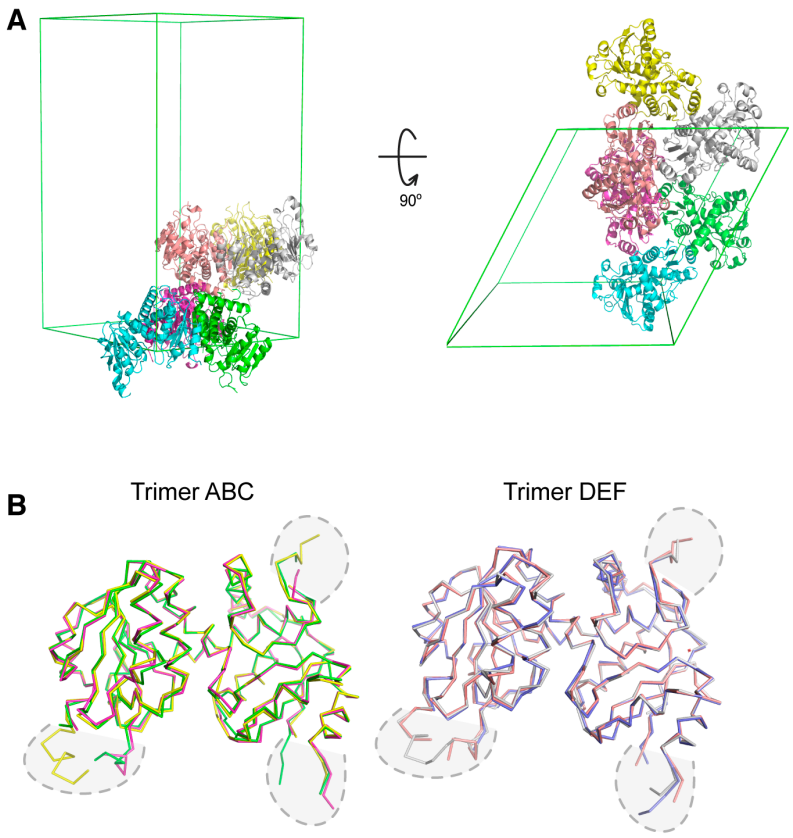

Figure 21. huATCase trimers lack a strict three-fold symmetry. (A) Representation of the monoclinic asymmetric unit with six protein subunits forming two trimers. **(B)** The superposition of the preliminary models of the subunits within the trimers suggested clear deviations from a strict three-fold molecular symmetry. Most important differences localized at flexible regions, encircled with a dashed line.

The largest differences between the subunits were found in loops that appeared to be more flexible. The different orientations of some side chains were also clearly noticeable. These discrepancies between the independent chains in the crystal, which might indicate conformational flexibility in solution, explained the problems during model refinement considering $P6_322$ as the space group, since the molecular three-fold axis is not a strict crystallographic axis. The models for huATCase-apo, huATCase-CP and huATCase-PALA were refined to final R/R_{free} values of 15.0/17.2, 22.1/25.1 and 15.2/20.8, respectively. The refinement statistics are as given in Table 6.

Table 4. Data collection and refinement statistics.

	huATCase-APO (PDB 5G1O)	**huATCase-CP (PDB 5G1P)**	**huATCase-PALA (PDB 5G1N)**
Data Collection			
Beamline	PXI, SLS	PXI, SLS	XALOC, ALBA
Wavelength (Å)	1.0	0.976	1.148
Space group	$P2_1$	$P2_1$	$P2_1$
Unit cell a,b,c (Å) α, β, γ (º)	82.8, 157.9, 82.7 90, 120, 90	83.7, 157.7, 83.5 90, 120, 90	83.0, 145.1, 83.2 90, 120, 90
Resolution (Å)	42.4-2.1 (2.15-2.10)[a]	50-3.20 (3.28-3.20)	50-2.1 (2.15-2.10)
R_{meas} (%)	11.6 (60.7)	17.1 (77.6)	6.8 (60.2)
I/σI	8.9 (2.4)	9.9 (2.3)	13.9 (2.5)
Completeness (%)	94.3 (63.9)	99.4 (99.9)	98.0 (97.9)
Redundancy	3.4 (2.9)	4.8 (4.9)	3.4 (3.4)
$CC_{1/2}$	98.9 (69.0)	99.0 (59.3)	99.8 (77.5)
Refinement			
Resolution (Å)	42.43-2.10	42.54-3.2	41.58-2.10
No. of reflections	95,831	29,277	92,437
R factor/R_{free} (%)	15.04/17.17	22.07/25.07	15.23/20.76
Rmsd Bond lengths (Å) Bond angles (º)	 0.008 1.284	 0.012 1.64	 0.007 1.268
No. of atoms Protein + ligand Water	 13,605 485	 13,098 95	 14,238 483
Ramachandran plot Favored (%) Allowed (%) Outliers (%)	 93.10 4.89 2.01	 88.4 8.24 3.4	 94.14 4.71 1.15

[a]Values in parenthesis correspond to the highest resolution shell.

8. Overall structure of huATCase

The molecular architecture of huATCase shares a common structural scaffold with bacterial ATCases and other enzymes of the transcarbamoylase superfamily. huATCase exhibits an equilateral triangular appearance, a dome-shaped 3-bladed propeller with active sites located at the concave or "bottom" face of the trimer (Figure 22). The radius from the center to the vertex of each subunit is roughly 50 Å, and the distance from the bottom to the top face is ~30 Å. Each subunit within the trimer is divided into two domains of similar size, nucleated by a central β-sheet of five parallel strands flanked by α-helices (Figure 22).

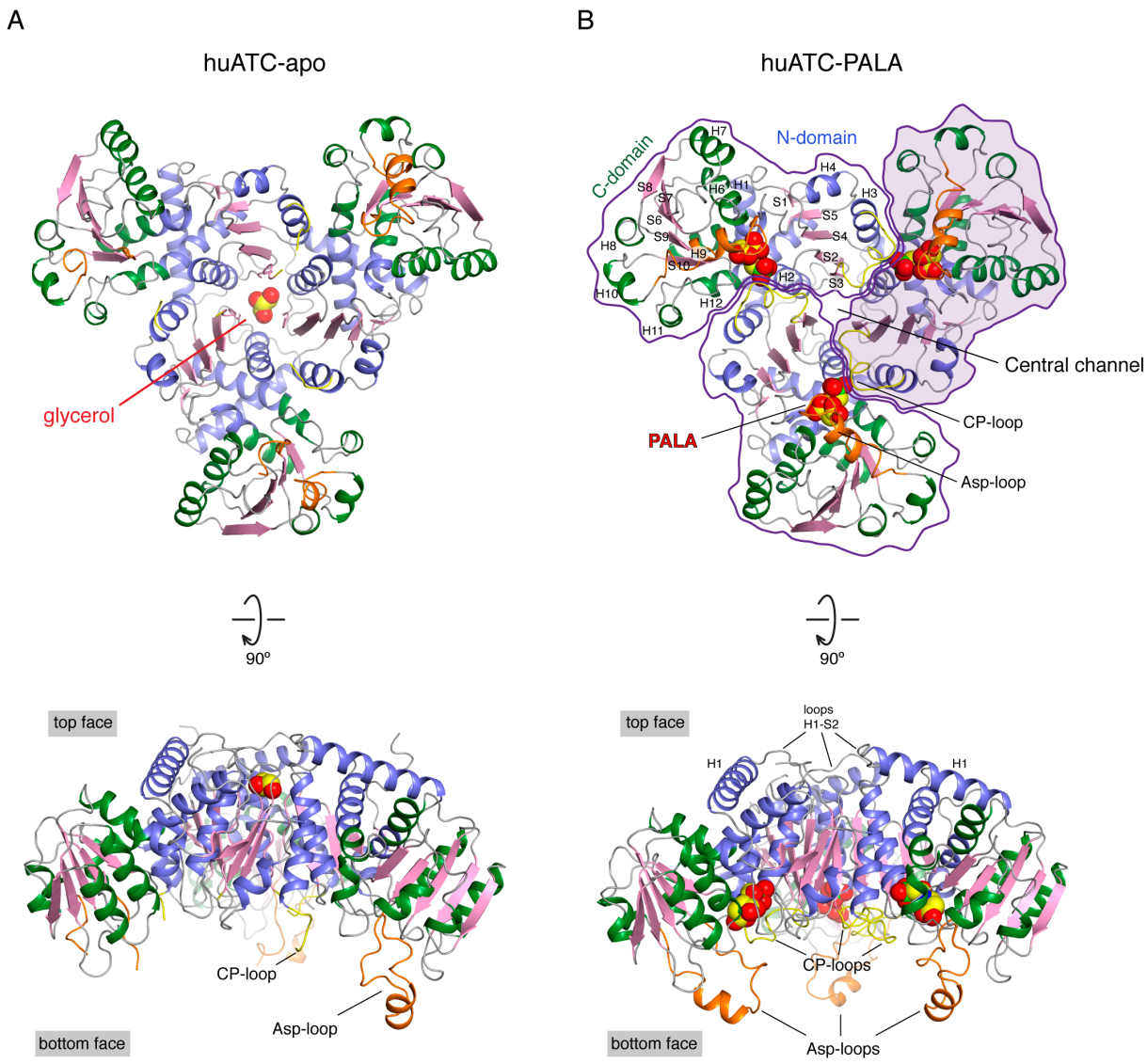

Figure 22. Overall structure of huATCase trimer. Cartoon representation of huATCase trimer in apo-form **(A)** or bound to PALA **(B)**. The trimers are shown in two perpendicular views. The central β-sheets are colored in pink, with the helixes of the N- and C-domains colored in blue and green, respectively. The CP-loop is shown in yellow and the Asp-loop is colored orange. The huATCase-APO structure has a glycerol molecule bound within the central tunnel. In **(B)** the area delimiting each subunit in the trimer is marked out with a purple line.

The two domains are interconnected by two α-helices –H5 and H12– running in opposite directions. The N-terminal domain (residues 1915-2056 and 2210-2225) is closer to the 3-fold axis, provides most of the contacts with the other subunits in the trimer and is responsible for the binding of CP (CP-domain) (Figures 21 and 22). The C-terminal domain (residues 2057–2209) occupies an external position at the vertices of the triangle, and holds the Asp binding site (Asp-domain).

The trimer has a central tunnel along the threefold axis, delimited by helix H2 and strand S3 from each subunit (Figure 22). The tunnel is 23 Å long and has a uniform radius of 3-3.5 Å along the axis, except at the top entrance where it is partially occluded by the loops connecting helices H1 and strands S2 (hereafter named loop H1-S2) and at the bottom face near the active sites, where the entrance is flanked by the CP-loops. The electron density was unambiguous for most of the huATCase-apo polypeptide chains, except for two disordered loops comprising residues 1994-2003 (hereafter named as CP-loop and equivalent to loop 80s in ecATCase) and residues 2148-2163 (named Asp-loop and equivalent to loop 200s in ecATCase) (Figure 22A). The Asp-loop appeared well-ordered only in one out of the six huATCase-apo subunits (subunit F) due to contacts with adjacent subunits in the crystal lattice. As already mentioned, soaking with CP damaged the crystals, but at least for one of them we collected a dataset. This huATCase-CP crystal diffracted at lower resolution (Table 6) and, although the overall quality of the density maps did not allow a precise tracing of some regions, a molecule of CP was assigned in the extra electron density observed within the active site of every subunit. In this structure, the CP-loops are well-ordered and interact with the CP bound in the active site of an adjacent subunit. On the other hand, the Asp-loops are flexibly disordered as in huATCase-apo crystals (Figure 22A). In turn, both the CP- and Asp-loops were clearly defined in huATCase-PALA structure, where a molecule of inhibitor is bound in the active site of every subunit (Figure 22B).

As mentioned, in huATCase-apo structure, the CP-loop and the Asp-loop are mostly disordered and thus the bottom entrance of the tunnel appears wide-open in the crystal structure. This might explain the presence of a glycerol molecule within the tunnel, next to residues G1988 and G1999 in the loop connecting helix H2 and strand S3 (Figure 22A). Contrary, in huATCase-CP and huATCase-PALA the rearrangement of the CP-loops causes the partial closure of the bottom entrance of the tunnel and the glycerol molecule is not observed despites being included in the crystallization and cryo-protectant solutions.

9. CP induces partial hinge-closure of the subunit and primes Asp binding

CP binds at the N-end of helix H2, with the phosphate group interacting with the main chain N atoms of residues T1974, R1975 and T1976, and with the side chains of T1976 and R2024 (Figures 23B and 24A). The carbamoyl moiety of CP interacts through the O atom with the side chains of T1976 and H2052, and through the N atom with the carbonyl O of

M2185 and with the side chain of Q2055. CP induces the positioning of the CP-loop from the adjacent subunit, which fits snuggly within the active site and binds to the CP through the side chain of S2000. The side chain of K2003 –at the tip of the CP-loop– is visible in one of the six subunits and points towards the active site but does not interact with CP. The repositioning of CP-loop is further stabilized by hydrogen bonds between the carbonyl O of T1998 and the N of S2000, both in the CP-loop, with the N atom of S1973 and the side chain of T1974 in the adjacent subunit.

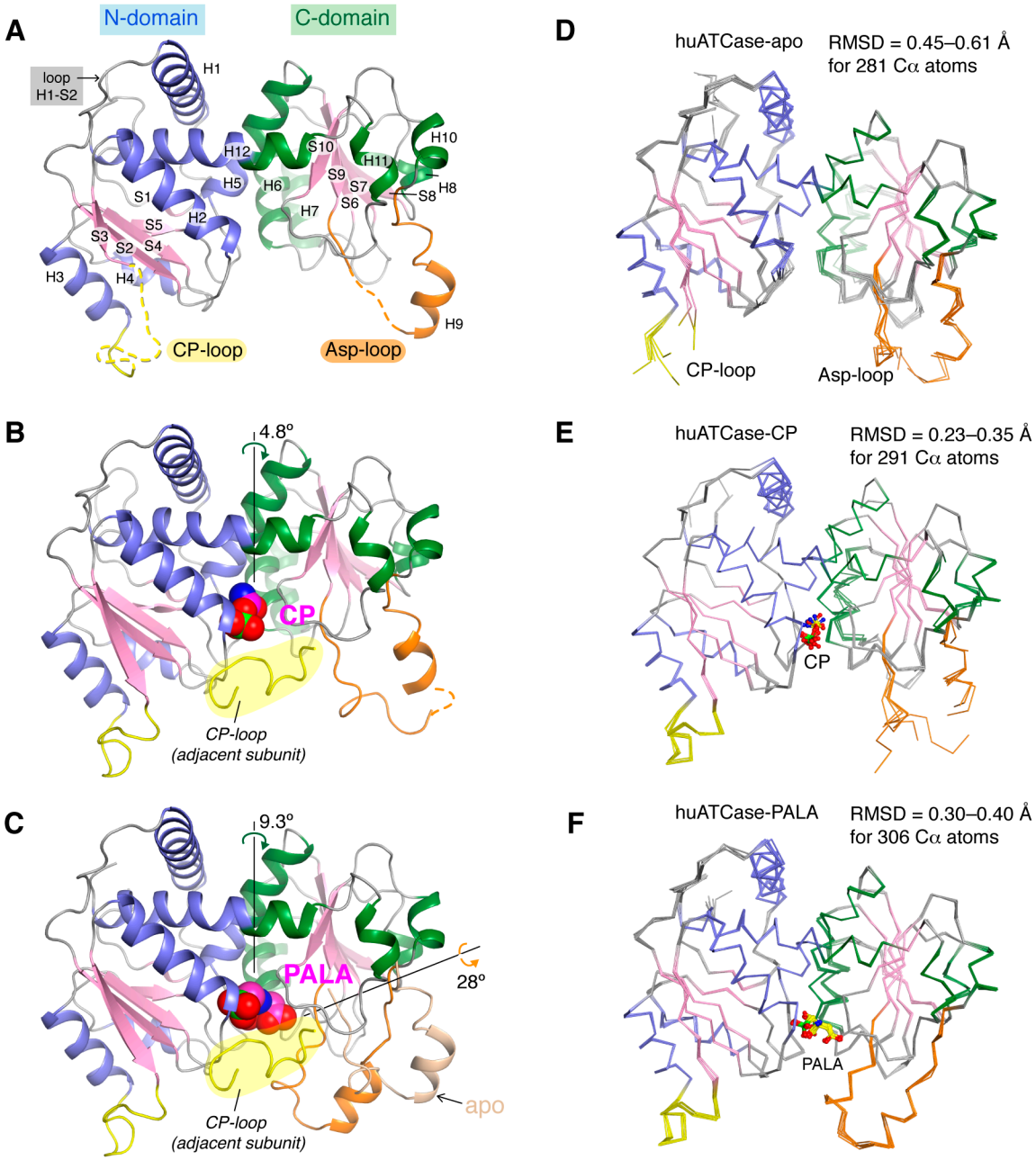

Figure 23. huATCase subunit and ligand-induced conformational movements. Cartoon representation of huATCase subunit free of ligands **(A)**, bound to CP **(B)** or in complex with PALA **(C)**. Ligands are shown in spheres and disordered regions are indicated in dashed trace. Secondary elements are labelled in **(A)** and coloured as in Fig. 22. Rotation axes for hinge-closure of the C-domain and for Asp-loop movement are indicated with black lines and curved arrows. **(D-F)** Cα trace superposition of the six subunits within the asymmetric unit. The root-mean-square-deviation (RMSD) for the position of the superposed Cα atoms is indicated.

In addition to the re-positioning of the CP-loop, the binding of CP results in an overall closure of the active site cleft, consisting on a 5º rotation of the C-domain around a hinge located at residues D2059 (helix H5) and A2207 (helix H12) in the two helices interconnecting the N- and C-domains (Figure 23B). Overall, the CP-induced movements alter the shape of the active site and drag the side chains of K2003 and R2146 towards a contiguous pocket adjacent to the CP site, which becomes primed for the binding of Asp.

10. PALA glues the N- and C-domains in a closed conformation

The simultaneous occupation of the CP and Asp binding sites by the molecule of PALA induces a more pronounced closure of the huATCase subunit compared to the CP-bound state (Figure 23C and 24B). The C-domain undergoes a rotation of 9º around the same hinge-axis described before, located at the helices interconnecting the N- and C-domains. In addition, compared to the observed position in subunit F in the huATCase-apo structure, the binding of PALA triggers a 28º rotation of the Aps-loop around a hinge located at residues 2145- 2147 and 2166- 2167 (Figure 23C). The Asp-loop harbours a conserved RxQxER sequence that confers specificity for Asp (Yuan et al., 1996). In this rigid body rotation, these residues undergo a ~10 Å displacement in order to interact with the inhibitor molecule (Figure 23B). The protein completely surrounds the molecule of PALA, excluding it from the solution, and interacts with every polar atom of the inhibitor (Figure 24B,C). The phosphonate group of PALA occupies the position of the phosphate group of CP, and interacts with the side chains of residues S1973, R1975, T1976 and R2024 and with the N atoms of R1975 and T1976 (Figure 24B).

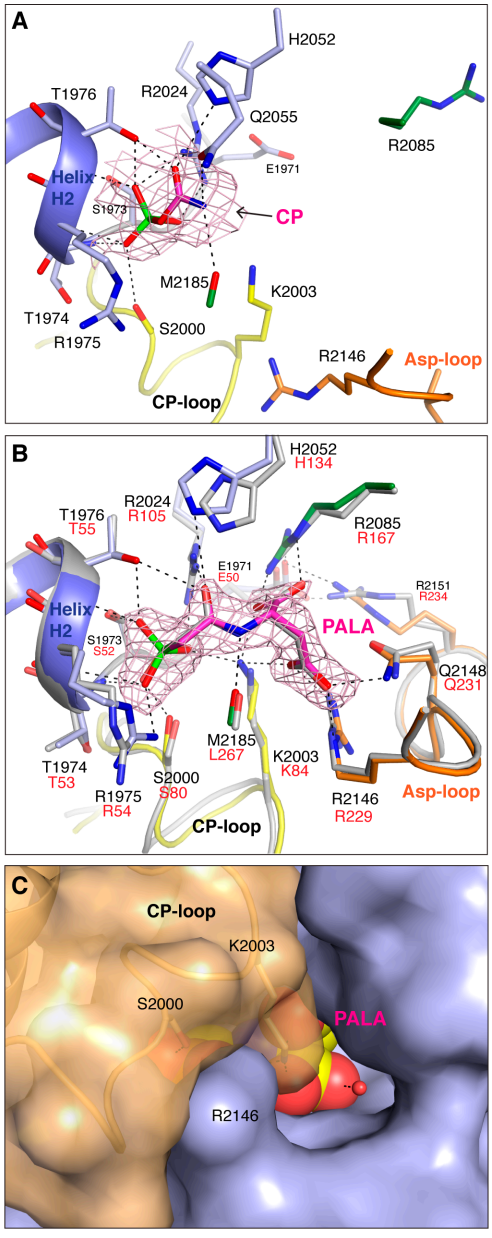

Figure 24. huATCase active site. (A) Active site of huATCase bound to CP. Ligands are colored with C atoms in magenta and residues are colored with C atoms in blue (N-domain), green (C-domain), orange (Asp-loop) or yellow (CP-loop from adjacent subunit). The electron density map for CP is shown. **(B)** Superimposition of PALA-bound active sites of huATCase and ecATCase-C3. huATCase is colored as in (B) and *E. coli* C atoms are colored in grey. **(C)** Surface representation of the huATCase subunit (colored in blue) with the repositioning of the CP-loop from the adjacent subunit (colored in brown). A molecule of PALA bound within the active site is shown as spheres. The surface is represented semitransparent and some of the protein elements are shown in ball-and-stick.

At the C-domain, the β-carboxylate of PALA interacts with the side chains of R2146 and Q2148 (in the Asp-loop), whereas the α-carboxylate complements the positive side chain of R2085. The carbonyl O of PALA is at hydrogen bond distance from the side chains of T1976, R2024 and H2052, and the N atom interacts with the carbonyl O of M2185. Moreover, the CP-loop from the adjacent subunit occupies the position observed in the huATCase-CP crystal and interacts with the phosphonate group of PALA through the side chain of S2000. The side chain of K2003 is well-defined in the huATCase-PALA structure and occupies a central position within the active site, with the side chain amino group at hydrogen bond distance from the α- and β-carboxylates, the N-atom and the phosphonate group of PALA.

Altogether, the interactions with PALA, which are summarized in Table 7, glue the N- and C-domains into a closed state further stabilized by a salt bridge between the side chains of R2151 (Asp-loop) and E1971 (loop preceding helix H2), by hydrogen bonding between R2151 (Asp-loop) and P2026 (loop S4-H4), and by extensive Van der Waals contacts between the Asp-loop and the CP-loop from the neighbour subunit.

Table 5. Interactions with PALA in the active site of huATCase. Asterisks show the residues from the adjacent subunit.

Residue	Residue atom	PALA atom	Distance A (Å)	Distance B (Å)	Distance C (Å)
S1973	OG	O3P	2.69	2.53	2.65
T1954	N	O2P	2.86	3.01	3.08
R1975	N	O2P	2.84	3.02	2.91
	NE	O2P	3.13	3.03	3.18
	NH2	O2P	3.15	2.89	2.57
T1976	N	O3P	2.87	3.03	2.73
	OG1	O1	3.18	3.24	3.18
	OG1	O3P	2.48	2.68	2.77
R2024	NH1	O1P	2.71	2.86	3.04
	NH2	O1	3.13	2.85	2.86
R2024	NH1	O3P	3.30	3.18	3.77
H2052	NE2	O1	3.16	2.91	3.30
R2085	NE	O2	2.91	2.55	2.75
	NE	O3	3.53	3.22	3.71
	NH2	O3	3.21	3.22	2.72
	NH2	O2	4.03	-	3.18
R2146	NE	O4	2.78	3.41	2.92
	NH2	O5	3.23	3.34	3.12
	NH2	O4	3.25	3.64	3.38
Q2148	NE2	O4	3.39	2.68	2.61
M2185	O	N2	3.02	3.01	2.77
S2000*	OG	O2P	2.95	3.10	2.94
	OG	O1P	3.28	3.04	3.24
K2003*	NZ	O5	3.11	2.80	2.95
	NZ	O3	2.88	2.98	2.65
	NZ	O1P	2.67	2.95	2.43

The PALA-induced motions are also coupled to a global change in the relative position of the subunits within the trimer (Figure 25). Looking at the huATCase trimer from the bottom face, the transition from the apo- to the PALA-state can be described as a 5º clock-wise rotation of the C-domains, with a 0.5–1 Å movement of the subunits towards the threefold axis. In a lateral view, the subunits come nearer at the bottom entrance of the central tunnel and separate by 1-2 Å at the top entrance delimited by the loops H1-S2. Upon PALA binding, there is also a displacement of 0.5-1 Å of helix H2 and strand S3, the two elements shaping the central tunnel, along the threefold axis and away from the active site (Figure 26).

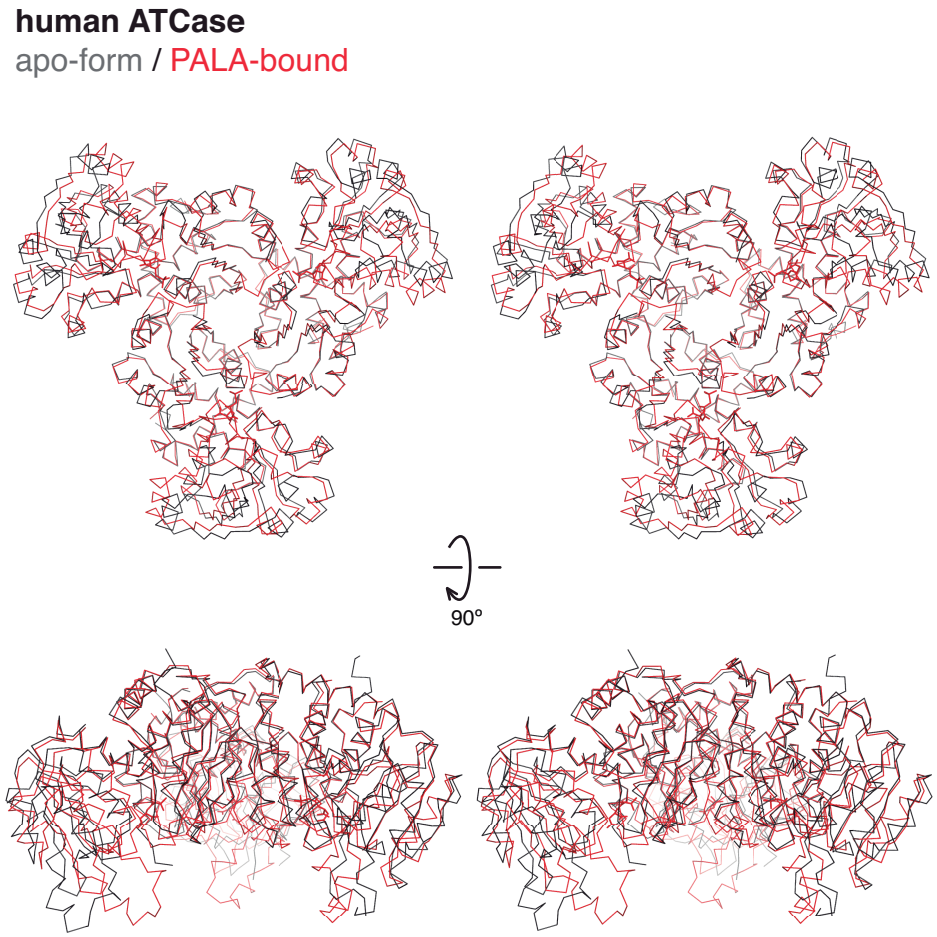

Figure 25. Conformational changes in huATCase induced by PALA. Stereoview of the superposition between huATCase trimer free (grey) and bound to PALA (red).

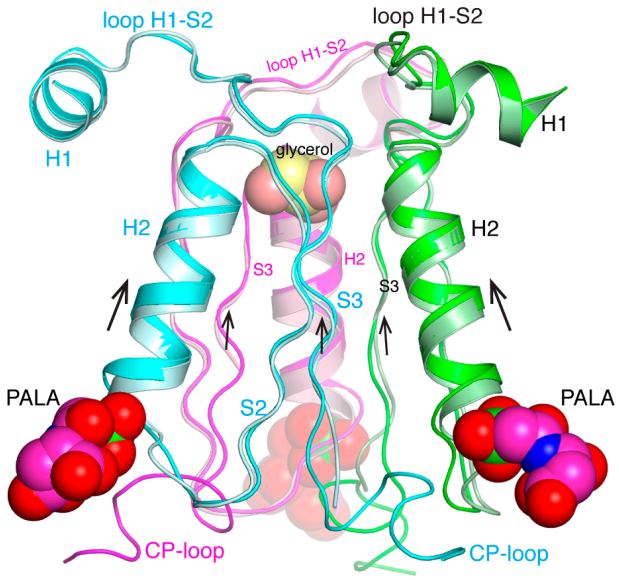

Figure 26. Ligand-induced conformational changes near the central tunnel. Cartoon representation of the superposition of huATCase free and bound to PALA, showing only the secondary elements shaping the central tunnel in perpendicular view to the threefold axis. The apo structure is shown in lighter colour tones. PALA and glycerol are represented as spheres. The arrows indicate the movement of helix H2 and strand S3 upon binding of PALA.

11. Comparison with *E. coli* ATCase reveals unique structural features in the ATCase domain of CAD

The huATCase structures confirm the predicted overall similarity with the catalytic trimers of ecATCase (Scully and Evans, 1991). Both proteins share a 44% sequence identity and the topology and length of the secondary elements are alike (Figure 27).

The subunits of human and *E. coli* ATCase free of ligands superimpose with an RMSD of 1.45 Å for 270 compared Cα's, whereas superposition of the PALA-bound conformations yield a RMSD of 1.06 Å for 296 Cα's (Figure 28). The two proteins are even more alike at the active sites (Figure 24B). The residues and interactions with CP or PALA are fully conserved with the exception of L267 in ecATCase-C_3 that is replaced by M2185 in human CAD (Figures 24B and 27). This residue contributes to the active site through the carbonyl oxygen and thus, the interactions with the ligands are the same in both proteins. In addition, the orientations of residues R1975 and H2052 are slightly different in the *E. coli* enzyme, but this could result from differences in the refinement of the models. Overall, the striking similarity between the two proteins and the virtually identical active sites strongly support that the catalytic mechanism proposed for ecATCase will fully apply to CAD.

Therefore, we asked what structural features could be responsible for the unique cooperative kinetic properties found in the human enzyme. The N- and C-termini of huATCase are four residues longer and three residues shorter, respectively, compared to ecATCase. Structural-guided sequence alignment with *E. coli* ATCase also revealed that huATCase presents two insertions and four deletions (Figure 27, highlighted in reddish).

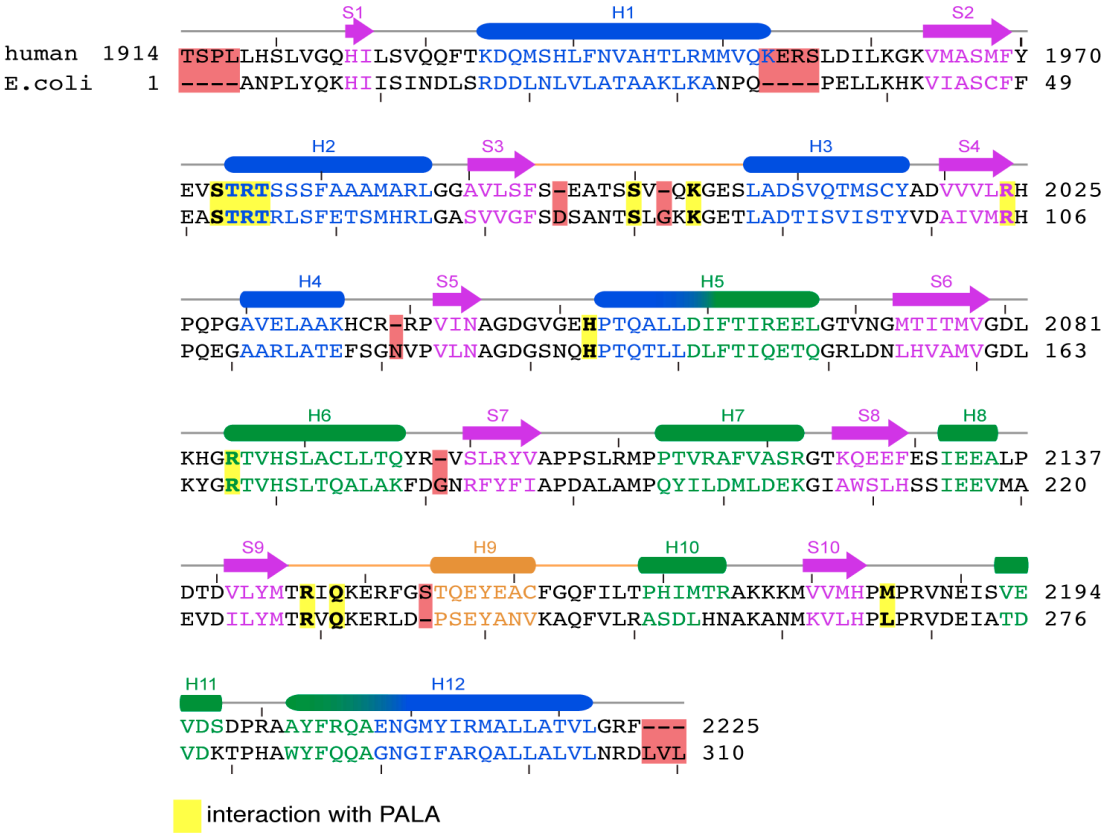

Figure 27. Sequence alignment of huATCase and *E. coli* ATCase catalytic subunit. Residues are coloured according to the secondary structure, which is drawn above the sequence. α-helices and β-strands are represented as cylinders and arrows, respectively. Residues interacting with PALA are shown in yellow background. The different insertions and deletions between the two sequences are highlighted with reddish rectangles.

Out of these, the most notable difference compared to ecATCase is the four-residue insertion at the C-end of helix H1 in huATCase. These residues fold as an additional turn at the C-end of helix H1 on the top face of the trimer (Figures 28 and 29). This helix is not only longer in huATCase, but it is also displaced 2-3 Å along its axis towards the adjacent subunit and tilted 7º towards the threefold axis compared to ecATCase. As a result, the loop H1-S2 is placed 8 Å closer to the center of the molecule and participates extensively in intersubunit contacts (Figures 29A). The first residue in the loop, E1954, which is invariant in CAD sequences but absent in bacterial ATCases, interacts with the side chains of residues R2224 and K1963 in the adjacent subunit. The loop residues D1958 and K1961 also interact with the same residues in the other subunits, forming a network of ion-pairs at the top entrance of the central tunnel (Figures 29B). These two residues are conserved in ecATCase (E37 and K40) but, based on the crystals structures, do not participate in intersubunit contacts (Figures 29B).

Another significant difference with *ec*ATCase is the two-residue deletion in the human CP-loop (Figures 27). As already mentioned, in the absence of ligands, this loop in huATCase is disordered (Figure 23A). In contrast, in ecATCase, this loop forms important

intersubunit interactions that involved residues D75 and N78 from different subunits even in absence of ligands (Figures 28 and 29). The interactions between *E. coli* subunits are reinforced by hydrogen bonding of R56 (helix H2) and G72 (strand S3). Thus, whereas in the absence of ligands ecATCase subunits are interrelated mainly through the interactions near the bottom entrance of the central tunnel, the huATCase trimer is hooked together through a network of ion-pairs on the top entrance of the tunnel, whereas the intersubunit contacts near the active sites are loose.

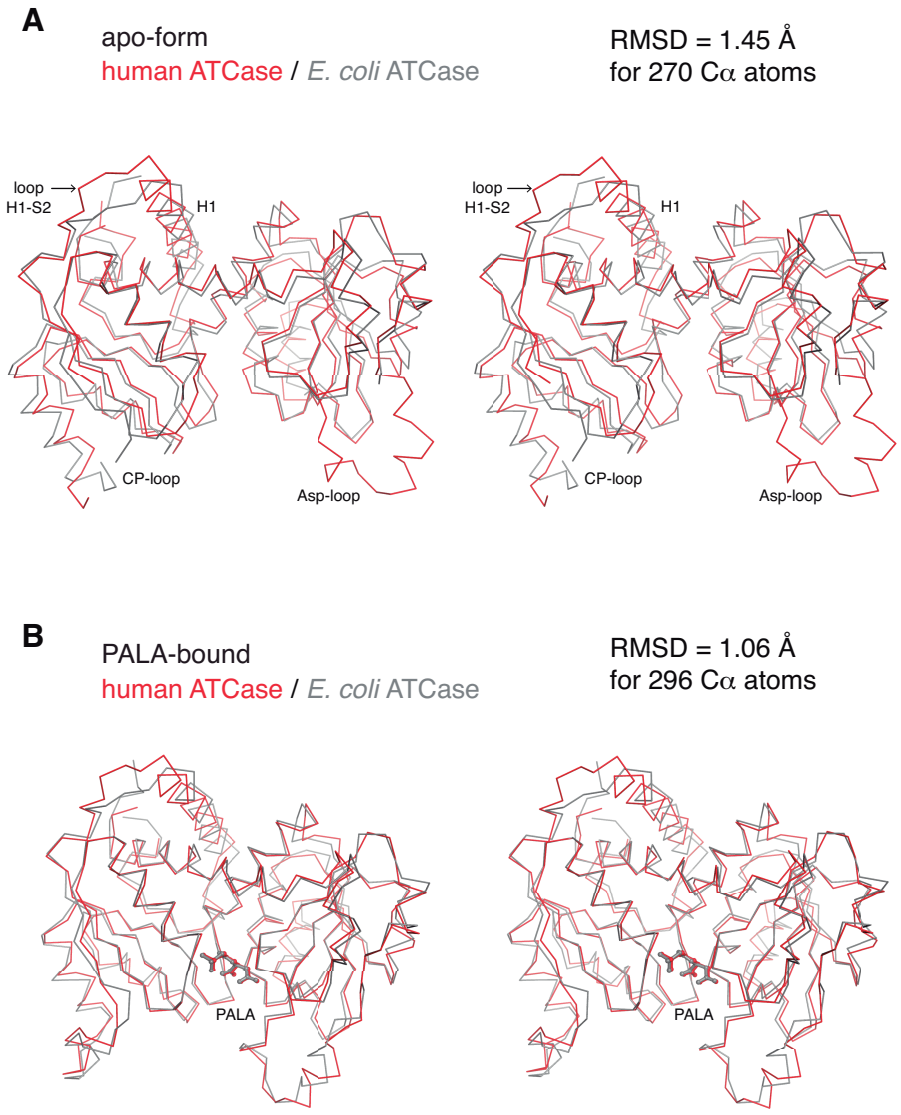

Figure 28. Structural comparison of huATCase subunit and ecATCase catalytic chain. Stereoview of the superposition of huATCase (in red) and ecATCase catalytic chain (in grey) free of substrates **(A)** or bound to PALA **(B)**. The subunits are shown in Cα trace and PALA molecules in ball-and-stick representation.

Upon PALA binding, the trimer shows a more compact conformation than ecATCase (Figures 25 and 28). The C-domains of huATCase are 1- 2 Å closer towards the adjacent subunits and helix H2, strand S3 and the CP-loop are 1-2 Å nearer the threefold axis. The smaller size of the residues in helix H2 facing the central tunnel, the shorter CP-loop and the lack of interactions at the bottom entrance of the tunnel may facilitate the approximation between the huATCase subunits (Figure 29C).

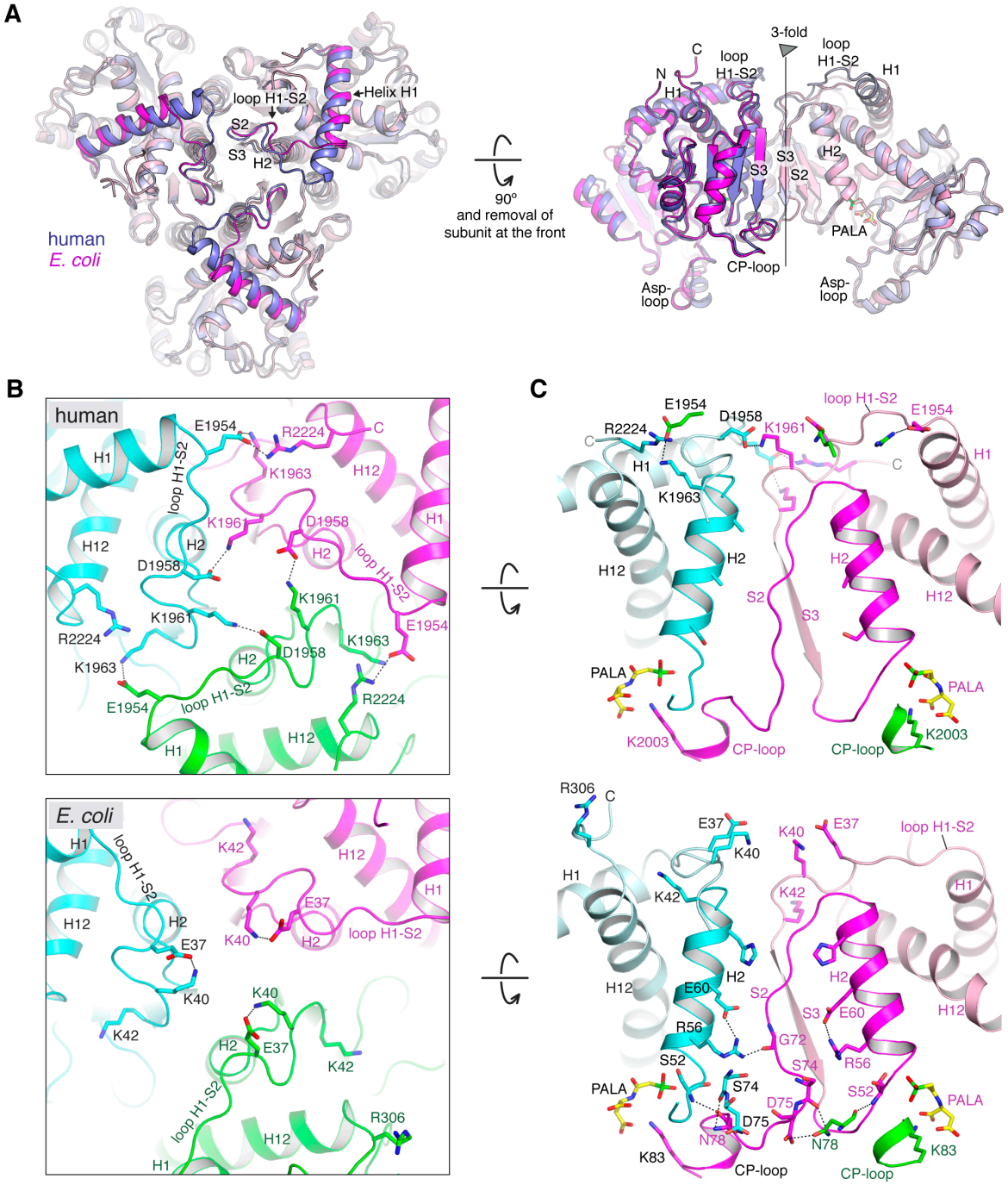

Figure 29. huATCase intersubunit contacts. (A) Superposition of huATCase trimer (purple) and ecATCase-C_3 (magenta) bound to PALA in two perpendicular views. Left, view along the threefold axis from the top face, with helix H1 and loop H1-S2 highlighted in a darker color. Right, lateral view with the front subunit removed and the two other two subunits colored in different tone. **(B, C)** Zoomed views showing the intersubunit contacts for huATCase (top) and ecATCase-C3 (bottom).

12. Loop H1-S2 interactions are key for huATCase stability and functioning

12.1. Purification of the H1-S2 mutants and determination of the oligomeric state

To test the significance of the loop H1-S2 interactions, we mutated residues E1954 and D1958 to alanine. Both mutants, huATCase-E1954A and huATCase-D1958A, were expressed and purified following the same protocol as for the wild-type protein (huATCase-WT), but the yield of huATCase-E1954A was 5-fold lower (Figure 30A).

We noticed that whereas in the final gel-filtration step the mutant D1958A eluted as a single peak similarly to huATCase-WT, the elution peak for huATCase-E1954A was double (Figure 30B). Based on the calibration of the size exclusion column with standards of known molecular weight, we estimated that this second peak could correspond to partially dissociated monomers or dimers of the protein.

SEC-MALS analysis further corroborated that similarly to the wild-type, the mutant D1958A forms trimers in solution at concentrations above 0.15 mg ml^{-1}, whereas the mutation E1954A causes a partial dissociation of the oligomers at similar concentrations (Figure 30C,D). We pooled the fractions corresponding to the trimer and re-injected the sample in the size-exclusion chromatography, this time at a concentration below 0.3 mg ml^{-1}. The elution profile of the second injection demonstrated that, at these concentrations, the oligomer is in association-dissociation equilibrium since a second elution peak was evident. The molecular mass measurement of this second elution peak was imprecise and did not allow us to discern if the species corresponds to a monomer, a dimer or a mixture of both.

Sedimentation velocity experiments showed two species (Figure 30E,F), one corresponding to the trimer and a second species with a sedimentation coefficient of 4.8 S and an estimated molecular mass of 40 kDa that matched the expected size of the monomer. The estimated dissociation constant of the E1954A and D1958A mutants is 4.5-fold and 2.7-fold higher than huATCase-WT, respectively. These results prove that the interactions at the top face of the trimer, absent in ecATCase, are key for the oligomerization of huATCase subunits.

12.2. Thermal stability

The mutations in the H1-S2 loop did not have a significant effect on the thermal stability of the protein. Thermal shift assays showed that, compared to wild-type (Figure 13), the mutations E1954A and D1958A do not significantly affect the T_M of the protein in the absence of ligands ($T_M^{E1954A,apo}$ = 48 ºC; $T_M^{D1958A,apo}$ = 51 ºC) (Figure 31A,B). Adding CP or PALA increased the T_M of both mutants to approximately 65 ºC and 76 ºC, respectively, which are lower values than those measured for the wild-type.

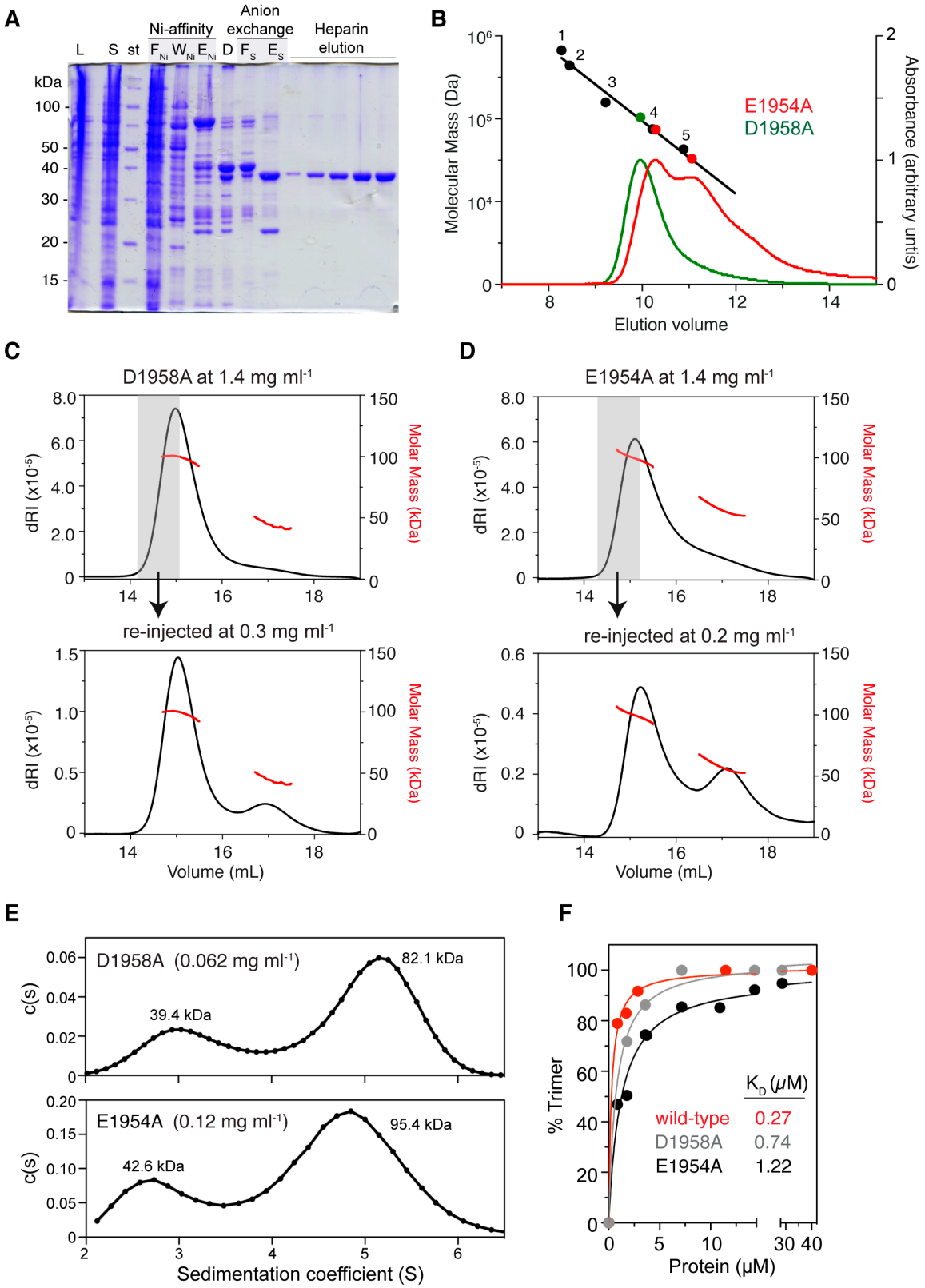

Figure 30. The loop H1-S2 mutants destabilize the trimer. (A) Purification of mutant E1954A monitored by SDS-PAGE. L, cell lysate; S, supernatant; F_{Ni}, W_{Ni} and $E_{Ni,}$ flowthrough, wash and elution from Ni-affinity column; D, dialyzed and digested with PreScission; F_S and E_S, flowthrough and elution from S column. **(B)** Semilogaritmic plot of the molecular mass versus the elution volume from a Superdex 75 column. The circles indicate the elution volume of proteins standards (in black; 1, thyroglobulin; 2, ferritin; 3, aldolase; 4, conalbumin; 5, ovalbumin) and of the E1954A (red) and D1958A (green) mutants. **(C, D)** SEC-MALS analysis. The differential refractive index and the molar mass are shown in black and red trace, respectively. In grey, peak fractions re-injected. **(E)** Sedimentation velocity assays and **(F)** estimation of the dissociation constants (K_D) for the mutants.

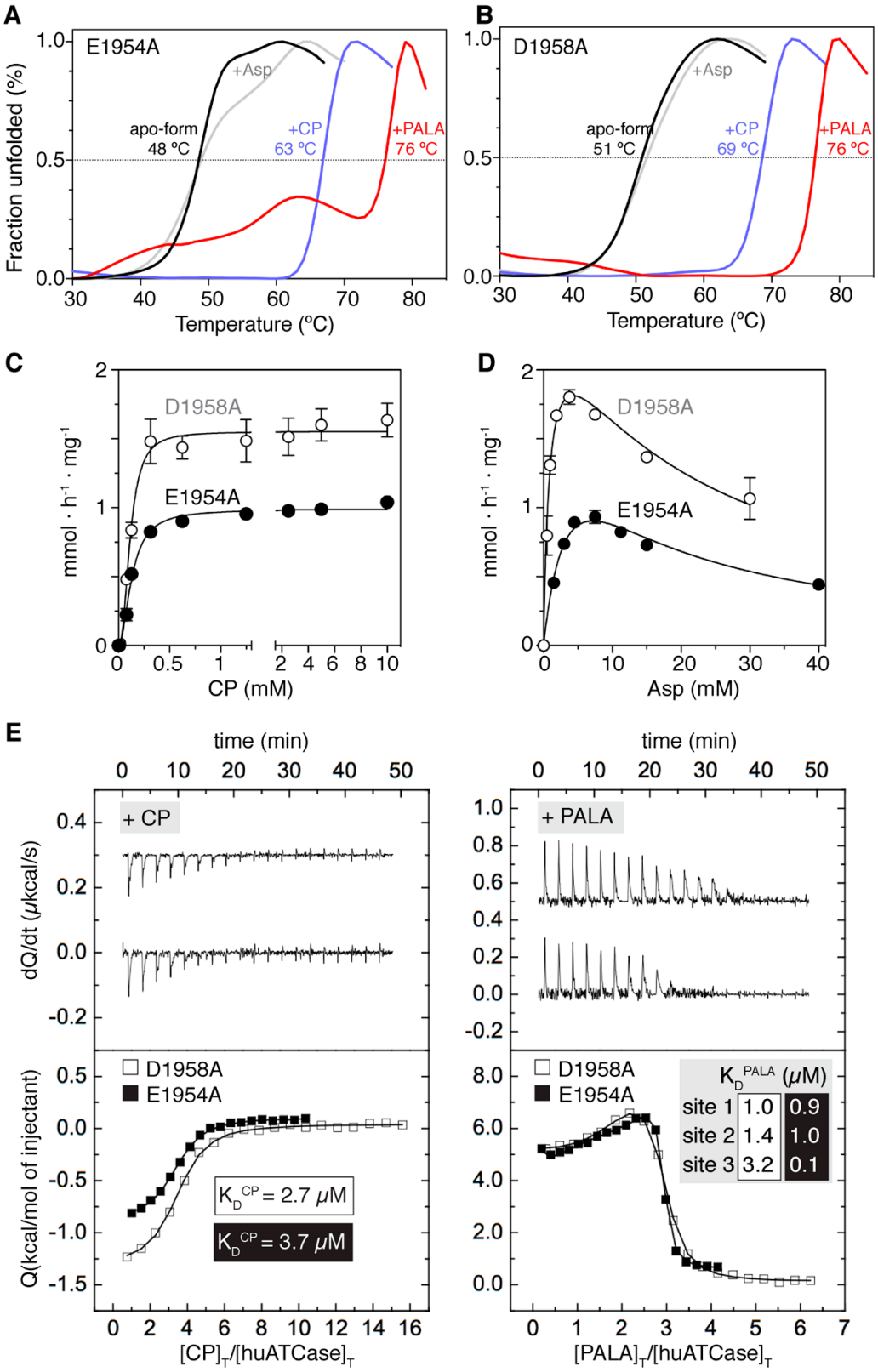

Figure 31. Stability, activity and ligand binding of the loop H1-S2 mutants. (A,B) Denaturing curves of E1954A **(A)** and D1958A **(B)** mutants measured in buffer Tris-acetate pH 8.3 in the absence or in the presence of ligands. **(C,D)** Substrate saturation curves at 10 mM Asp **(C)** or 5 mM CP **(D)** and varying the concentration of the other substrate. The kinetic parameters for mutant D1958A are V^{app}= 1.72 mmoles h^{-1} mg^{-1}, $[CP]_{0.5}$ = 0.11 mM, $[Asp]_{0.5}^{app}$ = 0.57 mM and n_H = 1.7- 2.4, whereas mutant E1954A shows V^{app}= 0.99 mmoles h^{-1} mg^{-1}, $[CP]_{0.5}^{app}$ = 0.13 mM, $[Asp]_{0.5}^{app}$ = 1.66 mM and n_H = 1.9. **(E)** Calorimetric titrations for the interaction of the mutants with CP (left) or PALA (right). The upper plots show the thermogram (thermal power as a function of time) and lower plots show the binding isotherm (normalized heat per injection as a function of the molar ratio ligand:protein).

12.3. Effects on enzyme activity and ligand binding

The loop H1-S2 mutations also altered the kinetic and the binding parameters of huATCase. Mutations D1958A and E1954A reduce the activity 2.5-fold and 4-fold, respectively (Figure 31C,D), likely due to the impaired protein oligomerization at the concentrations of enzyme used during the assays.

ITC experiments performed at protein concentrations (40 μM) at which both mutants form mostly trimers (Figure 30F) proved the binding of three CP molecules with similar affinities to wild-type (Figure 31E). However, the affinity and cooperativity for the binding of PALA are clearly affected. Both mutants exhibit similar dissociation constants for the first and second binding sites for PALA ($K_D^1 = K_D^2 \approx 1$ μM), which mean a ~50-fold lower affinity than the wild-type. Then, whereas D1958A mutant exhibits signs of negative cooperativity with a 3-fold decrease in affinity for the third site ($K_D^3 = 3.2$ μM), the E1954A mutant shows increased affinity of the third site ($K_D^3 = 0.1$ μM). These results demonstrate that disruption of the contacts through the loop H1-S2 somehow impacts on the activity and on the cooperativity between the active sites.

13. *In vitro* characterization of mutation R2024Q

The mutation R2024Q was recently described as the cause of the first case of a partial deficit of human CAD (Ng et al., 2015). The present structures show that R2024, similarly to the equivalent residues in ecATCase (R105) and in *B. subtilis* ATCase (R99), is a central element of the active site that directly interacts with CP (Figure 24A). Intriguingly, mutations of this arginine to alanine in the isolated catalytic ATCase trimer of *E. coli* (ecATCase-C3; mutation R105A) and in *B. subtillis* ATCase (R99A) were reported to confer positive cooperativity for Asp and to severely decrease the activity and the affinity for the substrates (Stebbins and Kantrowitz, 1992, Stebbins et al., 1989). To better understand the effect of this arginine and the effect of the R2024Q substitution in CAD, we produced and characterized the mutant huATCase-R2024Q.

The R2024Q mutant was expressed as the wild-type and the purification required small changes due to the reduced solubility of the protein at low salt concentrations (Figure 32A). The mutations does not affect the oligomeric state of the protein, as was demonstrated by SEC-MALS analysis (Figure 32B). The stability of the mutant in absence of ligands was also similar to wild-type (Figure 32C), indicating that the mutation does not affect protein folding. The increment in T_M upon addition of CP or PALA was 10 ºC lower compared with the wild-type and suggested a markedly reduced affinity for the ligands (Figure 32C,D). Indeed, ITC experiments proved that the affinity for CP was 7-fold lower than in the wild-type (K_D^{CP} = 42 μM), and we could only detect binding of PALA to one site per trimer and with very low affinity (K_D = 120 μM) (Figure 32E).

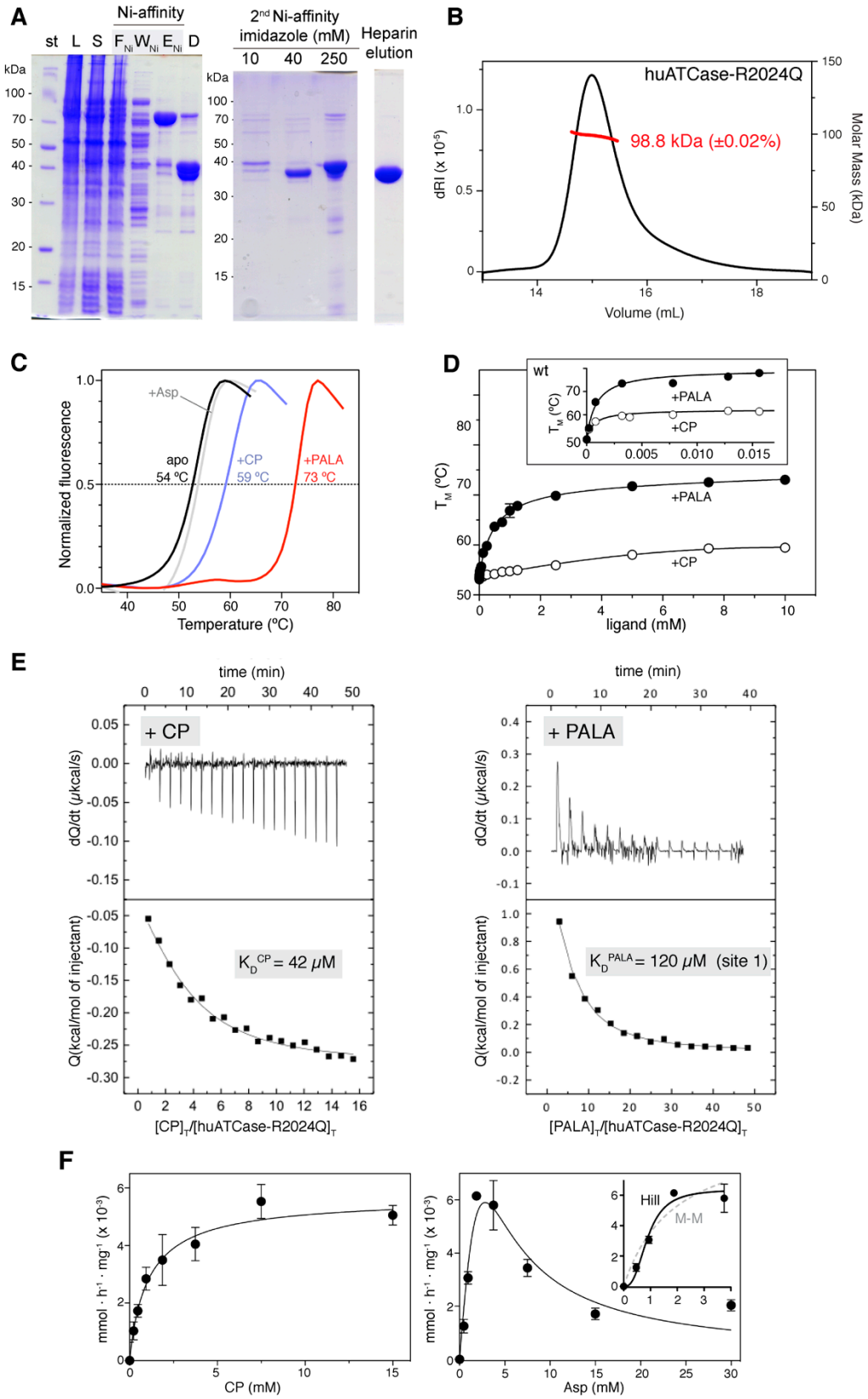

Figure 32. Characterization of R2024Q mutant. (A) Purification of huATCase-R2024Q monitored by SDS-PAGE. Fractions of the second Ni^{2+}-column indicate the elution at different concentrations of imidazole, with R2024Q eluting at 40 mM. **(B)** SEC-MALS analysis shows the formation of a stable R2024Q trimer. **(C)** Denaturing curves in the absence and presence of ligands. **(D)** Graphical comparison of the increment of T_M at varying concentrations of ligand for mutant R2024Q and for wild-type (inset). **(E)** CP and PALA binding to R2024Q mutant measured by ITC. **(F)** Substrate saturation curves at fix concentrations of Asp (10 mM, left panel) or CP (5mM, right panel). Inset, increased sigmoidictiy of the curve at low Asp concentration. Fitting to Michaelis-Menten or Hill equations are shown as grey dashed line or black lines, respectively.

The mutation virtually inactivates the enzyme, reducing the activity ~1000-fold (Figure 32F). The residual activity of R2024Q shows increased sigmoidicity, with a n_H = 2.7, a value that is similar to the ecATCase holoenzyme (n_H =2.6, (Mendes and Kantrowitz)), and also a more pronounced inhibition by Asp than the wild-type or the loop H1-S2 mutants (Figure 32F, right panel).

14. Isolation, mutagenesis and preliminary characterization of the bifunctional human DHOase–ATCase construct

14.1. Recombinant human DHOase-ATCase forms homohexamers in solution

It has been previously proposed that the 1.5 MDa CAD particle assembles by association of three polypeptide chains through their ATCase domains into a trimer, and then two of these trimers can further dimerize by interactions of the DHOase domains (Evans, 1986, Irvine et al., Carrey, 1995). Thus, CAD hexamer would result from a "dimer of trimers", with DHOase and ATCase playing a central scaffolding role in the structure of the particles. To test this hypothesis, we cloned a bifunctional construct encoding the human DHOase and ATCase domain (from M1456 to F2225; nucleotides 4365–6704). This construct also includes the 91 amino acid sequence linking both domains. This linker is rich in residues of proline (22%) and contains the residue S1859 that is phosphorylated by S6 kinase and PKA (Ben-Sahra et al., 2013, Robitaille et al., 2013, Carrey et al., 1985). The huDHOaseATCase construct was cloned into the pOPIN-F vector to introduce a cleavable N-terminal polyhistidine tag. The protein was expressed in human HEK293 cells (data not shown). This time, the addition of MBP was not required for protein solubility. Indeed, the presence of an MBP-tag at the N-terminus of the construct decreased the expression level (data now shown). Apparently, the presence of the DHOase domain is sufficient to allow the expression of the ATCase.

huDHOaseATCase was purified using three chromatographic steps: Ni^{2+}-affinity column followed by PreScission digestion of the His-tag, a second Ni^{2+}-affinity column and a size-exclusion chromatography (Figure 33). The purification yields ~3 mg of pure huDHOaseATCase per litre of HEK293 culture. The purified huDHOaseATCase migrated on a SDS-PAGE at the expected molecular weight of 84.70 KDa. The protein elutes from the size-exclusion chromatography as a major single peak at approximately 10 ml (Figure 32F), which corresponds to a complex of 560 KDa, in good agreement with the formation of a homo-hexamer in solution (theoretical molecular mass 508.1 kDa). A small shoulder in the void volume of the column indicates the presence of some aggregated material. A second shoulder at 12 ml also suggests partial dissociation into smaller oligomers. The analysis of the elution fractions by SDS-PAGE showed that the protein is highly pure and that all fractions in the main and secondary peaks correspond to the huDHOaseATCase (Figure 33E,F). The molecular mass of the oligomer was accurately measured by SEC-MALS using

a Superose 6 column. The protein eluted as a single peak with an average molecular mass of 469 kDa (± 0.15%) (Figure 34). However, the MALS measurement indicated that the sample was polydisperse.

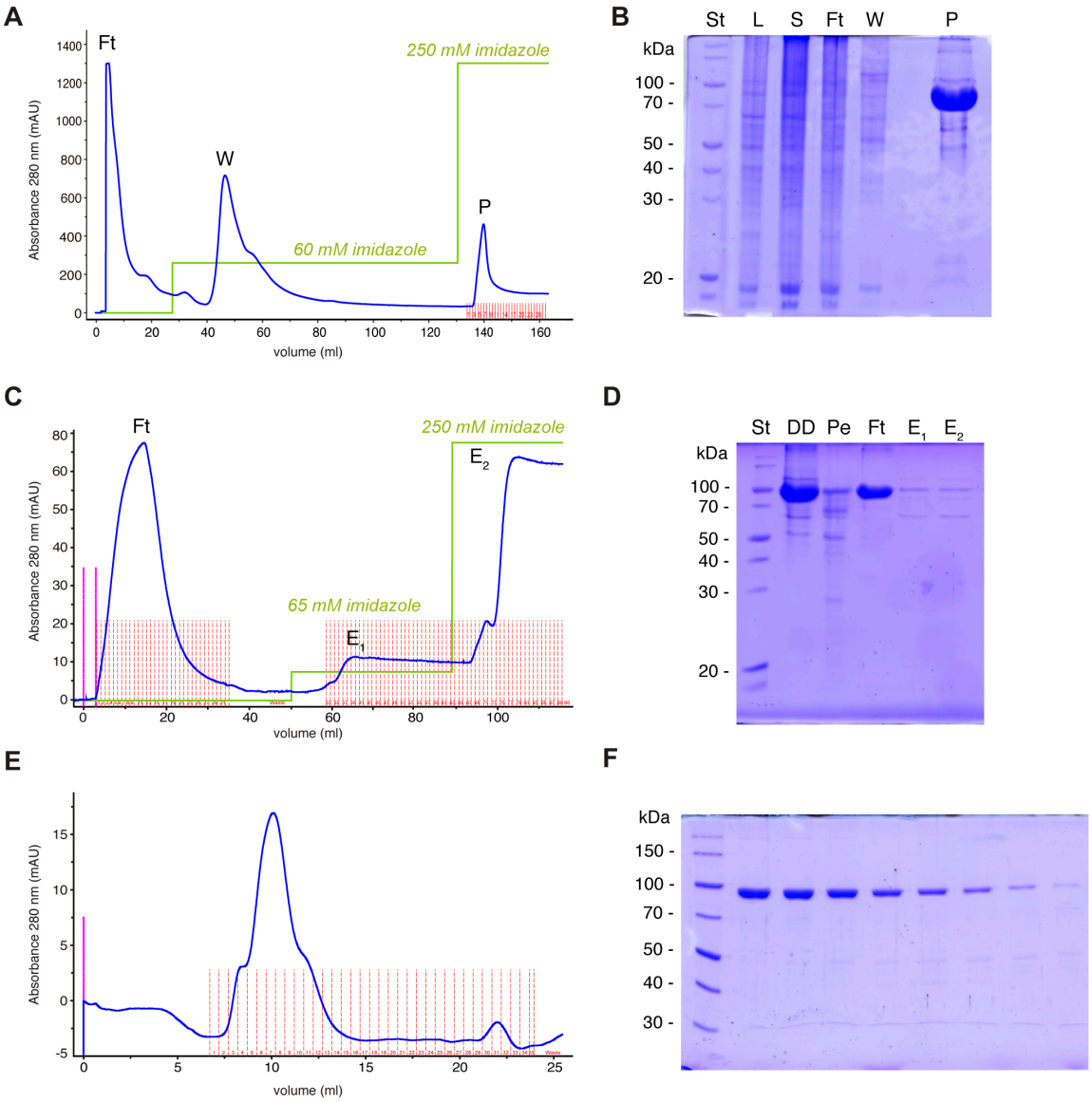

Figure 33. huDHOaseATCase purification. (A,B) Chromatogram and SDS-PAGE of Ni-affinity column. The blue line shows the elution profile monitored by UV absorption at 280 nm. L, cell lysate; S, clarified supernatant; F, flow through; W, wash with 60 mM imidazole; E, elution in 250 mM imidazole. **(C,D)** Following dialysis and digestion with PreScission (DD), the sample was loaded on a second Ni-affinity column**.** huDHOaseATCase does not bind the column (Ft) while the non-cleaved protein and contaminants remain attached to the column (E_1 and E_2). **(E,F)** Chromatogram of size-exclusion column (Superdex 200) and SDS-PAGE of the peak fractions.

14.2. Site-directed mutagenesis demonstrates the existence of a "dimer of trimers"

We introduced in the huDHOaseATCase construct mutations that are reported to prevent the oligomerization of the DHOase and ATCase domains. On the one hand, mutation M1601E was proved in our group to disrupt the dimerization of the isolated DHOase domain of human CAD (Grande-García et al., 2014). On the other hand, the mutation D2009A led to the dissociation of the isolated hamster ATCase trimers (Qiu and Davidson, 1998), mimicking the effect of the mutation D90 in ecATCase. In the context of full-length protein, the mutation D2009A was also shown to prevent the formation of hexamers (Qiu and Davidson, 2000). Both, the M1601E and D2009A huDHOaseATCase mutants were expressed and purified as the wild-type protein, although the recovery was approximately half than for the wild-type (1- 2 mg of pure protein per litre of HEK293 cells). Analysis by SEC-MALS demonstrated that the M1601E mutant eluted in a major peak with an average molecular weight of 271 kDa (± 0.33%), in good agreement with the formation of a trimer (theoretical mass 254.1 kDa) (Figure 34B). On the other hand, mutant D2009A eluted in a single peak with an average molecular weight of 182 kDa (± 0.50%), indicating the formation of a dimer in solution (theoretical mass 169.4 kDa) (Figure 34B).

We also mutated Ser1859 to glutamate or to aspartate to mimic the phosphorylation of this residue by the mTORC1 signalling pathway. It was reported that the phosphorylation of this residue promoted CAD oligomerization, a conclusion that was based on the formation of fluorescent foci and in changes in the sedimentation of the protein in centrifugation

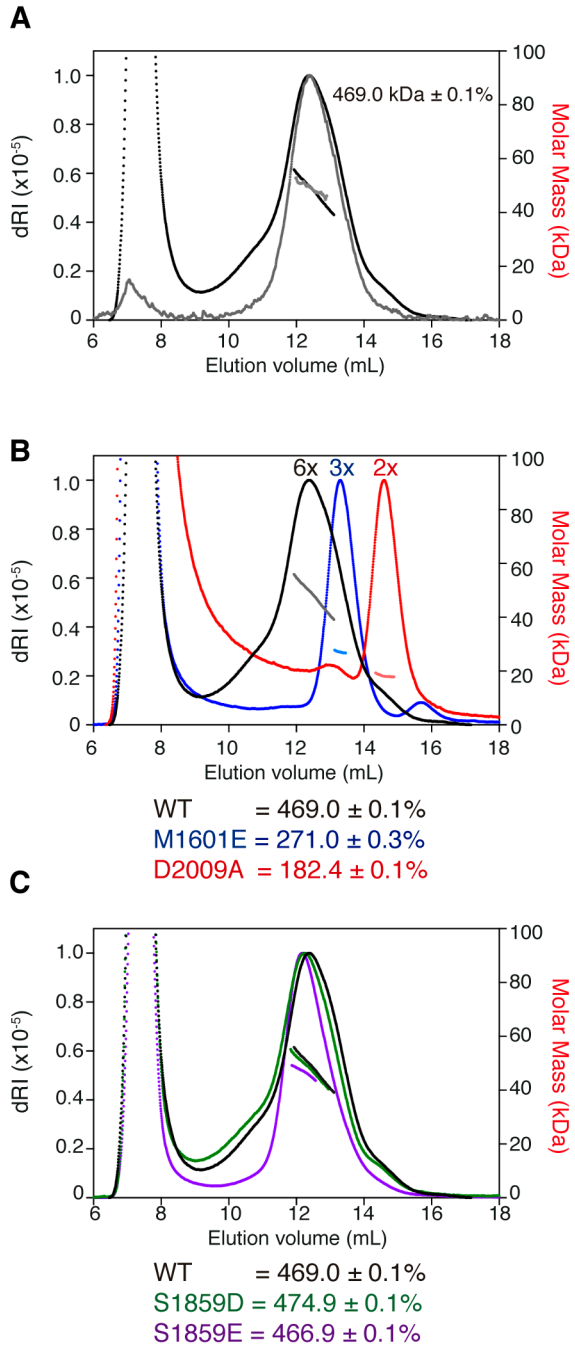

Figure 34. Determination of the oligomeric state of huDHOaseATCase. (A) SEC-MALS measurements of huDHOaseATCase-WT. **(B)** Superposition of the SEC-MALS measurements of huDHOaseATCase WT (black) and the mutants M1606E (blue) and D2009A (red). SEC-MALS analysis demonstrates that the oligomeric state is disrupted by the mutations. **(C)** Superposition of the measurements for WT (black), S1859D (green) and S1859E (purple).

gradients (Ben-Sahra et al., 2013, Robitaille et al., 2013). Once again, huDHOaseATCase mutants S1895E and S1895D were expressed and purified as wild-type. This time, we did not notice differences in the yield during purification compared with the wild-type. The analysis by SEC-MALS demonstrated that both mutants form homohexamers in solution with an average molecular weight of 271 kDa (± 0.1%). The mutant S1859D appeared to have more aggregated material than the wild-type. Contrary, the mutant S1859E eluted in a narrower and more symmetric peak than wild-type, suggesting a more compact and less polydisperse hexamer (Figure 34C).

DISCUSSION

1. First time production of the ATCase domain of human CAD

ATCase is the smallest of the four functional domains of CAD, but its role is central for the assembly of the 1.5 MDa hexameric particles. Pioneering work from J. Davidson's group demonstrated that a single-point mutation in the ATCase domain is sufficient to disrupt the oligomerization of hamster CAD (Qiu and Davidson, 2000). This result strengthened the assumption that CAD hexamers are formed by the dimerization of ATCase-mediated CAD trimers (Irvine et al., 1997). To get insight into the architecture of CAD, we aimed to isolate the ATCase domain for detailed structural characterization.

We succeeded in producing for the first time the recombinant ATCase domain of human CAD using both bacteria and human cells as heterologous expression systems. In the group, we routinely test the addition of different fusion partners or tags to the targeted proteins to improve their expression, solubility and purification. The ATCase domain of CAD is a clear example of the benefits of this strategy. We fused the N-terminus of huATCase to a hexahistine (His_6)-tag or to a dual His_6-maltose-binding protein (MBP) tag, but only the latter was efficiently expressed (Figure 10). These results were in agreement with a previous study from the group of Davidson reporting that the ATCase of hamster CAD was expressed soluble when fused to MBP (Qiu and Davidson, 1998). The precise mechanism by which MBP allows the expression of this domain of CAD has not been investigated. *E. coli* MBP is a large (43 kDa) stable and highly soluble protein, ranked as one of the best tags for enhancing the production and folding of passenger proteins (Kapust and Waugh, 1999, Fox et al., 2001, Costa et al., 2014). MBP has been reported to attract GroEL and to drive the tagged proteins into favourable chaperone environments for correct folding (Douette et al., 2005). Additionally, MBP presents a hydrophobic binding cleft, adjustable to accommodate different polypeptides, which can interact with the passenger protein (Kapust and Waugh, 1999, Fox et al., 2001). It is plausible that this intrinsic affinity endows MBP with a chaperone-like activity that assists in the folding of ATCase subunits by protecting exposed hydrophobic patches prior to their oligomerization.

The work of Davidson also demonstrated that the MBP-tag does not interfere with the activity of hamster ATCase (Qiu and Davidson, 1998). However, they did not test if the protein was stable upon removal of the MBP. We wanted to remove the MBP for crystallization purposes, and thus, a PreScission protease site was included in the protein construct to cleave off the tag. We proved that huATCase is soluble and stable in absence of the carrier protein, suggesting that MBP is only required for initial translation and folding. Unfortunately, all the advantages as a solubility enhancer turn into a curse when it comes to the purification of huATCase from the cleaved MBP, which becomes a major contaminant in the preparations. Additional chromatographic steps, such as a second Ni^{2+}-column, or amylose resins, proved to be inefficient for the binding of His_6-MBP. These difficulties made us suspect that, at high salt conditions favouring hydrophobic interactions, huATCase might

bind to the cleaved MBP reducing its affinity for the resins, as it has been described for other proteins (Fox et al., 2001). Luckily, at physiological salt concentrations, the separation of the two proteins is feasible using anion exchange chromatography. The huATCase sample freed from MBP showed only one other persistent but minor contaminant, identified as *E. coli* CAP protein.

We corroborated by SEC-MALS and AUC analysis that, similarly to other members of the transcarbamoylase family, huATCase forms trimers in solution (Figure 12 and 13). Although it is not the focus of this thesis, it is important to highlight that the observation that huATCase easily dissociates at concentrations ~0.01 mg/ml (K_D = 0.27 µM) has led us to re-design ongoing experiments aiming at visualizing full-length CAD by electron microscopy (EM). The preparation of grids for EM requires diluted samples at which we now know that favour the dissociation of the ATCase trimer. Now, dissociation of the CAD particles is prevented by chemical cross-linking which notably increases the homogeneity of the EM samples.

2. A picky protein leads to troublesome crystals

The first crystals of huATCase appeared in few hours, but four years later, the optimization of the crystallization conditions is still a pending matter. We do not understand yet how to control the conditions that lid from small thin and poorly diffracting plates to the growth of few thicker and well diffracting crystals. The other only factor that was found to make a difference in crystallization trials is protein freshness. Using the protein straight from the purification, without freezing or storage, was decisive to palliate excessive crystal nucleation, although it did not grant the reproducible growth of sizeable single crystals. Protein freezing was not only detrimental for crystallization but also caused partial protein precipitation and loss of activity. By assessing different freezing conditions, we concluded that the activity of the samples with no cryo-protectant or supplemented with 40% glycerol, decayed to 50% after three-weeks storage at -80 ºC. After thawing, part of the protein precipitated and the soluble fraction exhibited a decreased in V_M^{app} but similar $S_{0.5}$ values as the non-frozen samples, suggesting that part of the soluble protein was somehow damaged. The use of a 30% DMSO/ 5% glycerol had been reported in the past for the preservation of CAD (Coleman et al., 1977). This DMSO/glycerol mix did not prevent the precipitation of ~20 % of the protein after thawing, but conserved the specific activity of the soluble fraction. We suspect that although the sample is extensively centrifuged and filtered, the presence of a small fraction of precipitated or aggregated material could be the cause of excessive nucleation in the crystallization trials.

Since the refinement of crystallization conditions appeared fruitless, a brute-force screening of crystals was needed to find those of sufficient quality to collect complete datasets. In principle, the diffraction patterns were successfully indexed by invoking a single

lattice. The data was indexed and scaled in a primitive hexagonal lattice and space group $P6_322$. This indicated the presence of one huATCase subunit per asymmetric unit making a perfectly symmetric trimer in the crystal lattice. We did not notice any problem, except for the statistical distribution of intensities that deviated from what is expected for a reasonably untwinned data. Since this spacegroup admits no (pseudo)merohedral twin laws, it pointed to overmerging of pseudo-symmetric data or bad data quality. The molecular replacement with one *E. coli* ATCase catalytic subunit found a correct solution, but the model building and refinement was stalled at high values of R and R_{free}. Then, we re-scaled the data in lower symmetry space groups observing a marginal improvement of the statistics (Tables 4 and 5). The symmetry was eventually lowered to $P2_1$, and the refinement of two huATCase trimers per asymmetric unit led to acceptable R and R_{free} values. The crystal presents a "pseudosymmetry" problem, the molecular 3-fold axis is close to a true crystallographic symmetry, but the subunits within the molecule are indeed sufficiently different from each other. The non-crystallographic symmetry and the twinning operators combine to create an apparent 622 point symmetry. This problem was difficult to detect and could only be confirmed at a late stage during the refinement of the model. Luckily, twinning is not a fatal disease for non-expert crystallographers anymore and REFMAC5 is capable of working with multiple twinning operators, and to estimate and refine the twin fractions of the crystal between cycles of structural refinement (Murshudov et al., 2011), allowing us to determine the structure of huATCase.

3. Three snapshots along huATCase catalytic cycle

The unprecedented atomic view of the ATCase domain of CAD confirms the conservation of the overall fold and the virtually identical active site with bacterial homologues (Figures 24B and 28). The differences between the three huATCase structures, free of ligands, bound to CP and in complex with PALA, also indicate that similarly to other transcarbamoylases (Shi et al., 2015), the ATCase domain of CAD must undergo large conformational changes to complete the catalytic cycle (Figure 23).

In absence of ligands, huATCase shows an open fold with the CP- and Asp- loops flexibly disordered (Figure 22 and 23). The overall conformation resembles the ecATC-C_3 free of ligands but with fewer intersubunit contacts near the active sites. Then, as described for other ATCases, CP binds in first place and with high affinity to three independent and equivalent binding sites, causing a significant increase in protein stability by induced-positioning of the CP-loop over the active site of an adjacent subunit. The addition of Asp in absence of CP results in a negligible increase of protein stability, suggesting that there is no binding at the active site, or that the ligands do not trigger the conformational changes observed in the structures with CP (Figure 14). The positioning of the CP-loop fastens the loose intersubunit contacts at the bottom face of the trimer and together with partial approach

of the C-domain prepares the site for Asp. Binding of Asp must be concomitant with the closure of the active site to a conformation similar to the PALA-state, where the N- and C-domains compress the substrates in the correct orientations to react (Lipscomb and Kantrowitz, 2012). The closed huATCase-PALA structure shows three molecules of PALA engulfed by hinge-closure of their respective subunits and by rearrangement of the CP- and Asp-loops. Thermal shift assays have proven that this conformation is highly stable and cannot be mimic by the addition of CP and the inert Asp analogue succinate (Figure 14). Thus, the 10º hinge-closure of the C-domain and the ~30º rotation of the Asp-loop respond to a highly specific recognition of the ligand bound in the active site. Dissociation of PALA from this highly stable complex is impossible without reversing the conformational changes.

In the inhibition assays, where the enzyme is pre-incubated with PALA, we observed a substantial lag in catalytic activity before the linear formation of CAsp (Figure 18D). This lag is interpreted as a slow opening of the hinge and CP- and Asp-loops required for release of PALA and subsequent access by the CP and Asp. Indeed, the analysis of the kinetics of the interaction of PALA with ecATC-C_3, led to a mechanism involving a rapid binding of the bisubstrate ligand followed by a much slower isomerization of the complex (Cohen and Schachman, 1986). A similar behavior in huATCase, with a slow dissociation of PALA compared to substrate binding or catalytic turnover, would explain the behavior of PALA as a nearly irreversible inhibitor (Figure 18E).

4. Homotropic cooperativity in huATCase

Whereas the structural studies only provided unexciting similarities between huATCase and bacterial homologues, the characterization of kinetic and binding parameters pointed in a completely different and surprising direction.

The substrate saturation curves measured for huATCase did not obey Michaelis-Menten equation and formed slightly sigmoid curves, suggesting that the enzyme displays cooperativity between the active sites in the trimer (Figure 16). These results were quite puzzling since isolated bacterial ATCase trimers that shared strong structural similarity with the human enzyme were clearly reported as non-cooperative, a property that was reserved for the ecATCase holoenzyme (Porter et al., 1969, Gerhart and Holoubek, 1967). However, cooperativity was not the only plausible explanation for the sigmoidal kinetic behaviour. Activity is measured at protein concentrations (0.04 μM) 5-fold lower than K_D for the trimer, and although the dilution buffers contain 0.1 mg ml^{-1} BSA and this could have a crowding effect that favours oligomerization, it is possible that in the assay a fraction of the enzyme is in the inactive monomeric state. On the other hand, it is reported that CP –and PALA– promotes the association of ATCase trimers (Qiu and Davidson, 2000). Thus, the deviation of the Michaelis-Menten equation with abnormally reduced rates at low substrate concentrations could be related to the dissociation of the trimers. Nevertheless, the abrupt

flattening of the initial rates at higher substrate concentrations was another indication of a sigmoid and different from the rectangular hyperbola defined by the Michaelis-Menten equation.

We were only convinced of the cooperation between huATCase subunits by analysing the binding of PALA. The accurate measurement of different affinities for PALA in the three active sites of the trimer demonstrated without doubt the existence of homotropic cooperativity in the huATCase domain. Homotropic cooperativity arises when ligand-binding affinity of one active site is affected by the same ligand bound to a neighbour active site (Perutz, 1989). The evidence of such phenomenon in huATCase was clear, but a mechanistic explanation is troublesome since cooperativity shows a double face. On the one hand, binding of PALA decreases the affinity for the inhibitor at the vacant sites (negative cooperativity) but on the other hand, binding of substrates has the opposite effect and increases the affinity in the unliganded subunits (positive cooperativity).

Based on the structural data, we proposed a model to explain the complex kinetic behavior of huATCase (Figure 35). This model presupposes a progressive transformation of the trimer from an open to a closed state, where the crystal structures of huATCase free or bound to CP and in complex with PALA illustrate few of the possible conformations, with all the subunits in a similar state. Assuming the existence of intermediate states for which we do not have structural information, the negative cooperativity for PALA could be explained as follows (Figure 35). First, the empty trimer has a very high (nanomolar) affinity for PALA. Binding of the inhibitor to one subunit favors partial hinge-closure in the unliganded subunits and positioning of the CP-loop that might restrict the entrance of the phosphonate group of other PALA molecules. Binding of PALA to a second subunit prompts a further hinge-closure and rotation of the Asp-loop in the remaining unoccupied subunit that blocks the active site decreasing the affinity for PALA 100-fold.

A key question is whether the third active site unoccupied by PALA is still capable of efficient catalytic activity. Can the substrates bind to and react well in this vacant site, even if PALA cannot? We have not come up (yet) with a good experimental design to answer this question. Nevertheless, the finding that huATCase only binds PALA with high affinity in two out of the three active sites is surprisingly enough and relevant to understand –at least in part– the failure of PALA as an efficient anti-tumoral drug. The lack of therapeutic success is complicated, with major contributions from gene amplification (Kempe et al., 1976, Wahl et al., 1979), gut toxicity (Grem et al., 1988) and also from the ability of dietary uridine to bypass the PALA blockage through salvage pathways. However, at the molecular level, the possibility of a resistance mechanism based on the cooperativity between active sites should also be considered.

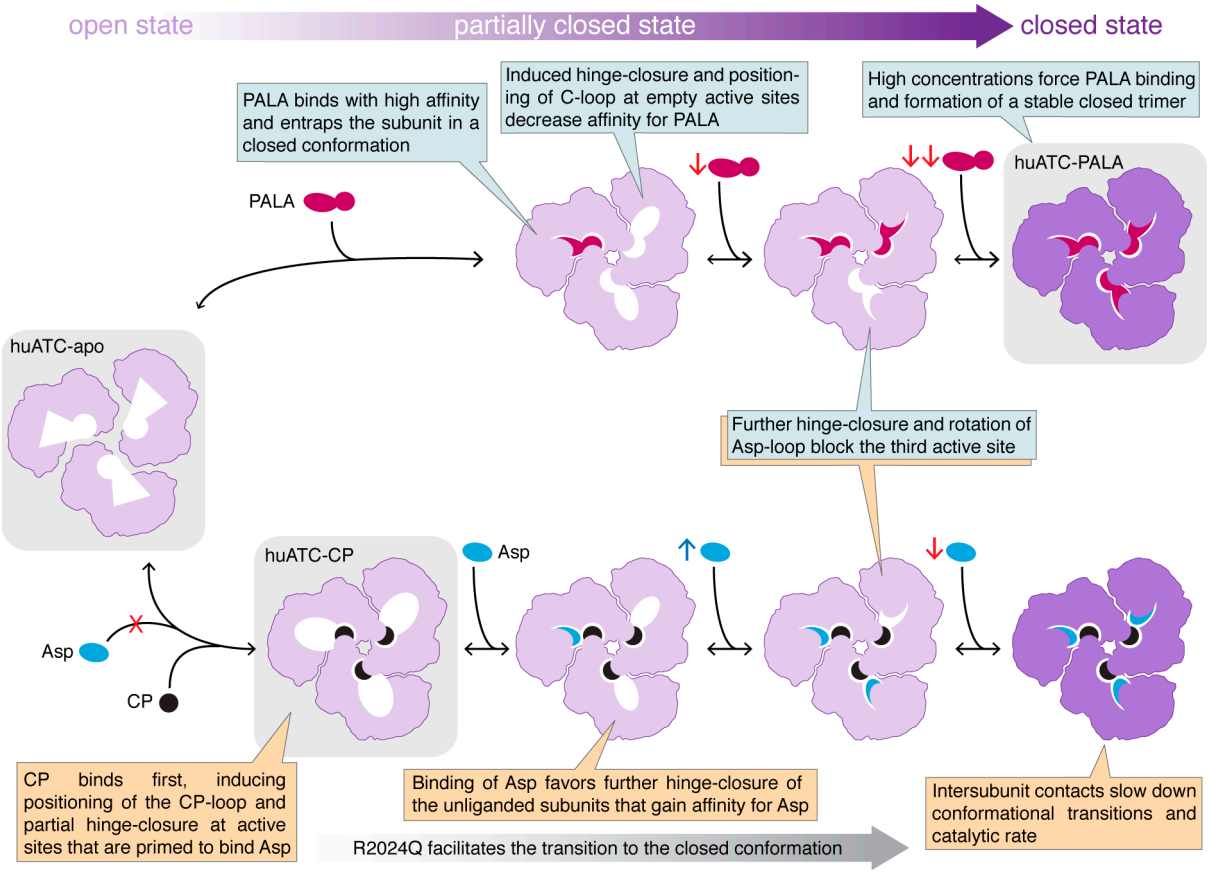

Figure 35. huATCase cooperativity model. Schematic representation of the transition of huATCase trimer from an open or partially open conformation (light violet) to a closed state (dark violet) with active sites saturated with PALA or with substrates. The grey background highlights huATCase states for which the structure has been determined. Blue and red arrows indicate respectively higher and lower affinity for the ligand.

Positive cooperativity, on the other hand, is manifested in the sigmoidicity of the kinetic curves (Figure 16). Since CP binds first and with equal affinity to the three active sites (Figures 19B), cooperativity must arise upon binding of Asp (England et al., 1994). We propose that binding of Asp to one subunit facilitates partial hinge-closure of the other CP-bound subunits (Figure 35). The positioning of the CP-loop and the approximation of the Asp-binding elements in the C-domain increase the affinity for Asp. Binding of Asp to a second subunit would predispose to the blockage of the third active site by further hinge-closure and rotation of the Asp-loop. This model implies that at intracellular concentrations of Asp (1– 4 mM) (Hansen and Emborg, 1994), only two subunits within huATCase trimer would be catalytically competent at a time. Nevertheless, it must be possible to force the binding of Asp to the three active sites, in the same manner as huATCase is crystallized with three molecules of PALA under high concentrations of the inhibitor.

5. The inhibition by Asp

Partial inhibition of the activity at high concentrations of the substrate to be carbamylated is a well-documented but poorly understood phenomenon in

transcarbamoylases (reviewed in (Shi et al., 2015)). As it has been shown previously for other ATCases, we observed that huATCase activity decreases at high Asp concentration (Figures 15B). It has been proposed that Asp could bind in a non-productive manner to the active site of ATCase, competing with CP, or to a second site in the enzyme (LiCata and Allewell, 1997, Pastra-Landis et al., 1978). During this study, we failed to detect the direct binding of Asp in absence of CP (Figure 14) and, as far as we know, other titration studies of the prokaryotic ATCase with Asp, succinate or malonate did not suggest stoichiometries greater than the number of active sites (Allewell and Licata, 1995). Furthermore, the crystal structures of prokaryotic ATCases obtained under high concentrations of ligands have not shown the aberrant positioning of Asp or of other analogues within the active site. Thus, at present, there is no direct evidence of the aberrant binding of Asp in the active site, although the possibility cannot be ruled out.

According to our model (Figure 35), huATCase activity has an upper limit of two active sites per trimer, that would correspond to the highest rate in the Asp saturation curves, at 5–7 mM of Asp (Figure 16B). However, we find plausible that the saturation of the three active sites with CP and Asp results in a conformation of the trimer similar to the PALA state, where increased intersubunit contacts could slow down conformational movements and reduce the catalytic turnover. Thus, substrate inhibition would be part of the allosteric and cooperative repertoire of ATCase. This idea was first introduced by N. Allewell based on analysis of kinetic data with different possible models (LiCata and Allewell, 1997), and finds now support in the current results with the human enzyme.

6. Which are the T- and R-states of huATCase?

To exhibit cooperativity, the enzyme must be in a conformational equilibrium between lower and higher affinity and/or activity states and to transit from one to another upon substrate binding (Monod et al., 1965). In the characterization of *E. coli* ATCase holoenzyme, the T-state was assigned to the conformation of the complex free of ligands (Kantrowitz and Lipscomb, 1988), whereas the structure of the holoenzyme complexed with PALA was assumed to represent the R-state (Kantrowitz and Lipscomb, 1988) (Figure 5). This was supported by the fact that substoichiometric amounts of PALA led to a complete conformational transition of the holoenzyme and to the activation of all the subunits in the complex (Macol 2001) (Kantrowitz, 2012). Lipscomb and colleagues beautifully described this transition between inactive and actives states of the holoenzyme, that involves a rotation and large separation between the catalytic trimers, with concomitant closure of the subunits over the molecules of PALA. However, when Schachman and colleagues solved the structure of the isolated ecATC-C_3 free and bound to PALA, realized of the similarity with the conformations observed in the holoenzyme, they challenged this mechanism of activation. They proposed that the activity of ecATCase is not dependent on the open or closed

conformations of the subunits, but on the fact that the movements required to catalyze the reactions are constraint in the structure of the holoenzyme free of ligands, whereas the separation of the catalytic timers upon PALA binding releases the movement of the loops. They also question that the structure of the PALA-liganded holoenzyme represented the R-state since it represents a blocked conformation of the subunits that impedes the free exchange of substrates and products in an out the active site.

Based on our data, we support the view of Schachman that the structure with three PALA molecules bound portrays a closed inhibited trimer unsuitable for interchanging substrates and products in and out the active site. As an analog of the true bisubstrate complex, the PALA-bound structure defines one of the necessary intermediates that must be accessible for the activated subunits to proceed along the catalytic cycle. It is probably that PALA freezes a conformation that would otherwise only occur transiently, when the enzyme stabilizes the transition state. Thus, in our view, the PALA-structure does not represent the active state of the protein. It is only one of the many conformations that huATCase needs to go through in order to catalyze the reaction. On the other hand, we also argue that the structure of the enzyme free of ligands is not a good representation of the T-state neither, since this conformation is only posed to bind CP and shows no affinity for Asp.

Since we disagree with the association of the free and PALA-bound structures with the T- and R-states, we asked if any of the conformations proposed in our model (Figure 35) could account for the equilibrium between low and high affinity/activity conformations required for cooperativity. As it has been shown previously, cooperativity only arises upon binding of Asp (England et al., 1994) and implies the equilibrium of different states with lower or higher affinity for this substrate. The enzyme is only primed to bind Asp if CP loads first in the active site. Thus, we propose that the structure of huATCase in complex with CP would be a good representation of the enzyme in the T-state. Then, we propose that the trimer loaded with CP plus one Asp bound to one of the subunits would represent better the R-state, since the two subunits that do not have Asp would exhibit increased affinity for this substrate. Thus, cooperativity in huATCase would be restricted to the conformational equilibrium between huATCase-CP and bound to one Asp molecule (Figure 35). A structure partially saturated with PALA or with CP and Asp (or a inert mimic) would be excellent to validate the model. Unfortunately, neither this work or the structural characterization of bacterial ATCases have shown a catalytic trimer partially saturated with ligands.

This could explains why we failed to detect activation by adding small concentrations of PALA or succinate (Figure 17). In *E. coli* ATCase holoenzyme, the inhibition of one subunit by the binding of succinate (or PALA) favors the T>R conversion, and thus, the activation of the five other subunits in the complex (Stebbins et al., 1989). In contrast, in huATCase, the inactivation of one subunit upon binding of succinate or PALA, would result in the activation of only two subunits. Furthermore, according to our model (Figure 35), if one of the two

activated subunits binds Asp, the remaining subunit will not be catalytically competent due to the induced blockage of the active site. In this manner, the activation would be much lower than the ecATCase explaining that we could not detect it by the colorimetric assays.

7. Structural bases of cooperativity

The isolated ecATC-C_3 or other bacterial ATCases that are only catalytic trimers are repeatedly described as being non-cooperative (Gerhart and Holoubek, 1967, Porter et al., 1969, Stebbins et al., 1989). However, landmark studies by Howard Schachman reported that only two ecATC-C_3 subunits bound PALA with high affinity, while affinity of the third site was much lower (Newell et al., 1989). This discrepancy with the identical binding sites observed in the crystal structures was attributed to protein damage during isolation and thus, it might have not received enough attention. We do not find reasons to suspect that huATCase is damaged since, at least for CP, the protein shows three identical and equivalent binding sites (Figure 19B). Thus, it is tempting to reinterpret these data also as a negative cooperativity effect in the binding of PALA to ecATCase-C_3. Nevertheless, since ecATCase-C_3 kinetic curves are not sigmoidal (Stebbins et al., 1989), the cooperativity effects –if any– are diminished compared to huATCase. Although we were convinced of the tight structural similarity in the overall fold and virtually identical active sites between huATCase and ecATCase-C_3, the different kinetic behaviour presupposes structural differences between the subunits, and in the way they interact. In this manner, finding these differences called for a second and careful comparison of the two structures.

The most distinctive feature of huATCase is the tight interconnection through ion-pair networks involving the loops H1-S2 (Figure 29A,B). These loops are located at the top end of the trimer, surrounding the 3-fold axis and 30 Å away from the active sites. We proved that mutations in this loop compromise the oligomerization of huATCase and decrease the affinity and cooperativity for PALA. However, ecATCase lacks the interactions through this loop, and thus, the stability of the *E. coli* trimer must depend on the tighter intersubunit contacts at the bottom entrance of the central tunnel (Figure 29C). Then, we asked how the huATCase loop H1-S2 can affect the functioning of the active site, located on the opposite face of the trimer, and participate in the propagation of conformational changes between subunits. We observed that closure of the active site upon PALA binding leads to a 0.6-1 Å movement of helix H2 and strand S3 along the threefold axis and away from the active site (Figure 26). The movement of these two elements shaping the central tunnel directly affects the position of the loop H1-S2 and the connection between subunits. Thus, we propose that changes of the loop H1-S2 in the unliganded chains could initiate the sequence of events leading to hinge-closure in the absence of ligands. Additionally, the more compact conformation of huATCase compared to ecATCase-C_3 in the PALA state (Figure 25) also suggests that communication of conformational changes could be enhanced by the proximity between

subunits. We demonstrate that single-point mutations in this loop compromise the oligomerization of huATCase (Figure 30B,C,D). Residue E1954 is fully conserved in all the CAD sequences that we have found available and it appears to be central for the oligomerization of the trimers. Mutation D1958A also reduces the stability of the huATCase trimer but the effect is less pronounced. Interestingly, upon addition of CP or PALA, the mutants show similar melting temperatures as wild-type, indicating that binding of the ligands favours the oligomerization of the protein (Figure 31A,B).

Even when the subunits of ecATCase-C_3 and *B. subtilis* ATCase appear to act independently, it has been shown that a single mutation of an arginine in the active site is sufficient to confer positive cooperativity for Asp (Stebbins and Kantrowitz, 1992, Stebbins et al., 1989). The huATCase structures show that the position and interactions of the equivalent residue, R2024, are fully conserved with bacterial ATCases (Figure 24B). Remarkably, R2024Q has been identified as the causing mutation of the first and only known partial deficit of human CAD (Ng et al., 2015). As expected from the position in the active site (Figure 24A,B) and from the effect of the mutation in bacterial ATCases, R2024Q strongly reduces the activity and decreases the affinity for CP and PALA. Most interestingly, R2024Q mutant increases the cooperativity and causes an earlier onset in the inhibition by Asp (Figure 33F). This indicates that R2024Q mutation facilitates the transition to a closed trimer, and further supports that increased intersubunit contacts may slow down conformational movements and enzyme turnover.

We propose that R2024 –or the equivalent arginine in bacterial ATCases– is part of a conformational switch operated by the binding of ligands to the active site (Figure 36). In the "open position", the switch arginine points outside of the empty active site, impeding hinge-closure of the C-domain by steric hindrance and by blocking the interactions with E1971.

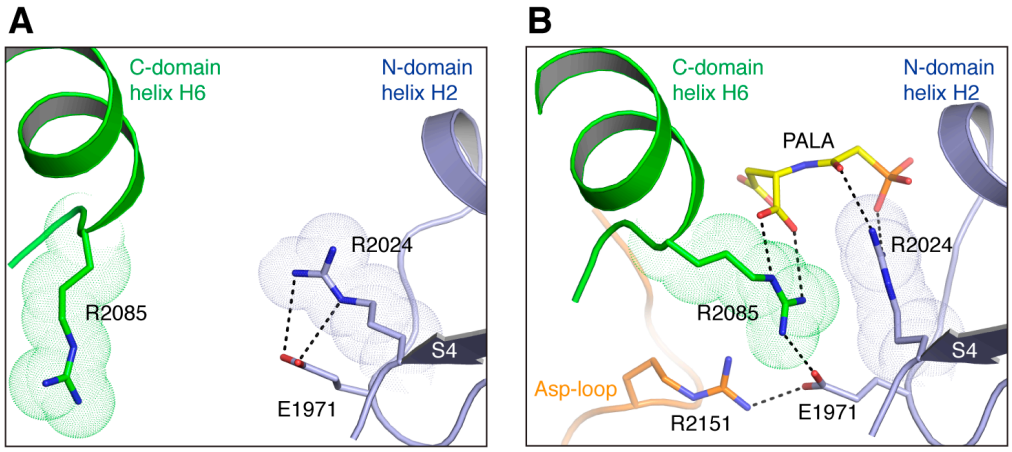

Figure 36. R2024 conformational switch. Detailed view of R2024 interactions in the active site of huATCase free (left) or bound to PALA (right).

Binding of ligands switches the arginine to a "closed position" that allows the approach of the C-domain, the correct positioning of R2085 to interact with Asp and the stabilization of the Asp-loop over the active site. Mutation of the switch arginine increases cooperativity by releasing the interactions that restrain the subunits in a low-affinity state, and displacing the equilibrium towards a conformation with higher-affinity for Asp.

8. ATCase domain in the context of the complete CAD

This work, and previous works carried out in the group, have resulted in the detailed characterization of the DHOase (Grande-García et al., 2014) and ATCase domains of human CAD. Despite the acquired knowledge about the nuts and bolts of these two domains, we lack information on how they assemble together with GLNase and CPSase into hexameric particles (Coleman et al., 1977, Lee et al., 1985). There are no many ways in which the CAD polypeptides can arrange into a hexamer preserving the symmetries of the DHOase dimers and ATCase trimers (Lee et al., 1985, Evans, 1986, Irvine et al., 1997). We support a model (Carrey, 1995) in which three CAD polypeptides associate through their ATCase domains forming trimers. Two of these trimers could further dimerize between their respective DHOases making hexamers. Thus, the ATCase and DHOase domains would have a central role in the formation of CAD "dimer of trimers" (Figure 37). In this study, we provide data supporting this model. We proved by SEC-MALS that the recombinant huDHOaseATCase construct behaves as a hexamer in solution (Figure 34). The sample shows certain polydispersity, which could indicate high flexibility of the construct in solution. This flexibility can be expected taking into account the lengthy 91-residue linker connecting DHOase and ATCase domains. We further corroborated our model by site-directed mutagenesis. The mutant M1601E is a trimer in solution, since the mutation hinders the formation of DHOase dimers (Figure 34). In turn, mutant D2009A is a dimer in solution, as the mutation destabilizes the ATCase trimer. Thus, our results indicate that the DHOase and ATCase domains form the core scaffold of the CAD particles.

Based on single-particle electron-microscopy (EM) data obtained by other colleagues in the group, we have developed a model for the architecture of the CAD particle (Figure 37). By using negative staining techniques, we have obtained 2D-averaged images of preferred particle orientations of the DHOase-ATCase construct. We combined this low resolution information with the high-resolution structures of the human DHOase and ATCase domains to build a hypothetical model of the particle. Similar to the models first proposed by Elizabeth Carrey and envisioned by David Evans, we placed the two huATCase trimers at the edges of the particle, joined together through the interaction with three huDHOase dimers.

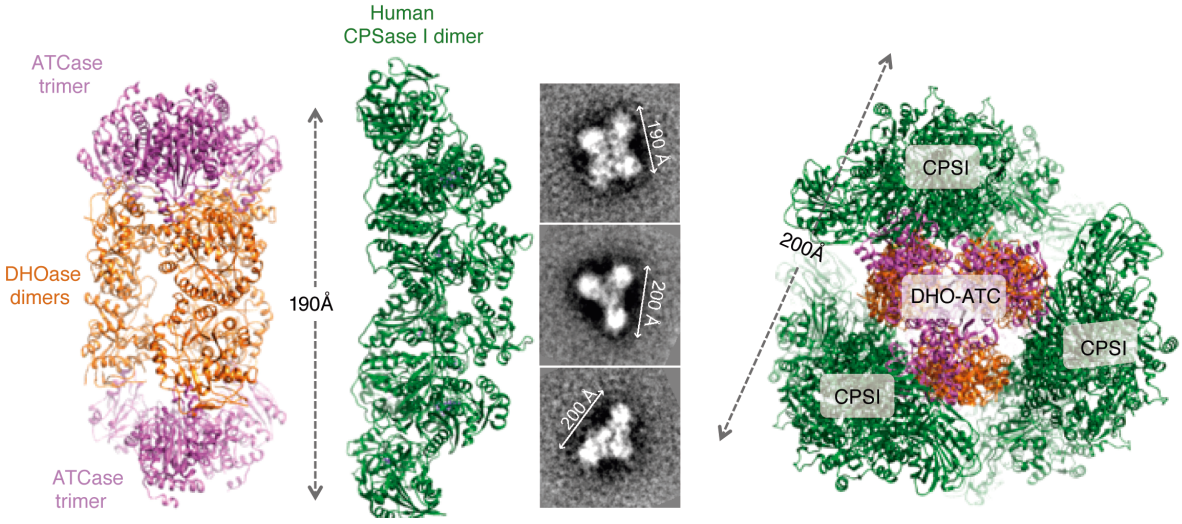

Figure 37. Hypothetical model of human CAD based in EM and crystallization techniques. DHOase and ATCase crystal structures are built together based in the MALS and mutagenesis data. CPSase structure shown corresponds to the CPSase I structure solved in the group of Vicente Rubio.

This arrangement resembles the disposition of the ATCase trimers in ecATCase holoenzyme or in the *A. aeolicus* ATCase complex with DHOase, although the human particle is compact and more elongated, expanding approximately 190 Å. We think that these two domains assemble to form the structural core of the particle and that the GLNase and CPSase domains are ensemble in a second layer around this central element. Future EM experiments with full-length CAD should finally reveal the architecture of the protein at high resolution. At the moment we have been unable to isolate the GLNase and CPSase domains for structural characterization, but we were able to produce full-length CAD protein from other organisms (hamster and fungi) and the EM characterization is underway. For the time being, we can only build and work on plausible models. Doing so, we were very intrigued by the interesting crystal structure of the human CPSase I determined by the group of Vicente Rubio (de Cima et al., 2015). This structure showed the formation of CPSase I dimers within the crystal with approximately the same length as the DHOase-ATCase particles. Assuming that the CPSase domain of CAD could make dimers similar to the mitochondrial counterpart, we propose that six GLNase-CPSase domains could wrap the central DHOase-ATCase core forming three dimers. This would result in a compact architecture of the particle where the orientation of the domains would favor the communication between active sites. The huATCase active sites would be facing the DHOases, while the exit for the long tunnel traveling across the GLN-CPSase domains would communicate with the central segment of the particle. This arrangement of the CAD subunits seems to fit well with very preliminary negative staining EM data collected for the CAD-like complex from *Chaetomium thermophilum* (work in progress) but there is still a long run ahead to reach the final atomic view (or views) of this amazing protein.

Regarding the enzymology of the complex, we have not measured yet the enzymatic properties of the DHOase-ATCase construct. We hypothesize that since the catalytic rate of ATCase is directly related to hinge-movement of the C-domain and to the flexibility of the CP- and Asp-loops, it is plausible that these movements are conditioned by the association of ATCase with other CAD domains, resulting in a decreased activity of the complex. Indeed, the ATCase activity of hamster CAD is ~10-fold higher in the isolated trimer than within the full-length protein (Qiu and Davidson, 1998, Qiu and Davidson, 2000).

Moreover, the potential ATCase activity within CAD is further limited by the supply of CP that is synthetized by CPSase at ~50-fold lower rate (Coleman et al., 1977, Qiu and Davidson, 2000). This difference in rates between CPSase and ATCase could be advantageous in ensuring that the unstable molecules of CP –produced by CPSase– are immediately captured and used by the ATCase.

This predicted slower-performance of huATCase within the CAD complex, conditioned by interaction with other domains, might explain the clinical effects of mutation R2024Q. Although this mutation reduces the activity of the isolated domain 1000-fold it does not completely abolish CAD function since huATCase in the CAD complex is not required to work at full rate. Furthermore, it has been described that there is a channeling of endogenous CP between the CPSase and ATCase active sites (Christopherson and Jones, 1980). This might help to explain how CAD may overcome the detrimental effect of the mutation R2024Q on the affinity for this substrate.

These thoughts are relevant to understand that inhibiting the active site of ATCase domain might not be the most effective strategy to compromise the activity of CAD. Targeting the CPSase domain, which catalyzes the limiting step of the overall reaction and is the point of allosteric control, could be much more effective for drug design.

We also speculate on the possibility that similarly to the *E. coli* ATCase holoenzyme, where the association through regulatory dimers confers cooperativity to the catalytic subunits, the association of two huATCase trimers within the CAD complex might enhance the cooperativity effects described in this study. It will be very interesting to assay which is the cooperativity of ATCase within the DHOaseATCase construct. If the cooperativity turns out to be increased, this would have implications for the fine-tuning of CAD activity to small variations in the intracellular concentration of Asp. On the other hand, higher cooperativity between subunits could also result in larger impediments for PALA binding in CAD particles partially saturated with the inhibitor. Further insight into the functioning and inhibition of the ATCase domain awaits the detailed characterization of the DHOaseATCase complex and on the entire CAD particle.

CONCLUSIONS

1. We have produced the ATCase domain of human CAD as an independent and catalytically active recombinant protein.
2. The isolated huATCase domain forms trimers in solution for which the dissociation constant has been estimated.
3. We have determined the crystal structures of huATCase in the apo-form, bound to CP and in complex with the inhibitor PALA
4. The crystal structures corroborate the predicted overall similarity with bacterial ATCase catalytic trimers.
5. huATCase is a homotrimer, where the subunits undergo large conformational changes upon binding of ligands to the active sites.
6. Contrary to the isolated bacterial ATCase catalytic trimers, huATCase shows positive cooperative for substrate binding and negative cooperativity for PALA binding.
7. We combined structural analysis and mutagenesis to identify the distinctive structural elements responsible for the communication and cooperativity among huATCase subunits.
8. We have produced and characterized the huATCase mutant R2024Q, described to cause the first human disease associated with a CAD-deficit. We propose that residue R2024 is part of a conformational switch controlling the conformational changes of the subunits upon ligand binding.
9. We produced a bifunctional construct harboring the ATCase and DHOase domains of human CAD and proved that they form the hexameric core scaffold of the CAD particles.

CONCLUSIONES

1. Hemos producido, de manera recombinante, el dominio ATCasa de la CAD humana (huATCasa), que se comporta como una proteína independiente y es catalíticamente activa.
2. El dominio huATCasa aislado forma trímeros en solución, para el cual hemos calculado su constante de disociación.
3. Hemos determinado las estructuras cristalográficas de huATCasa en la forma apo, asociada a CP y en complejo con el inhibidor PALA.
4. Las estructuras cristalográficas corroboran la similud con los trímeros catalíticos de las ATCasas bacterianas.
5. Las subunidades de huATCasa sufren grandes cambios conformacionales al unirse los ligandos al sitio activo.
6. Al contrario que los trímeros catalíticos de las ATCasas bacterianas, huATCasa muestra cooperatividad positiva para la unión de los ligandos y cooperatividad negativa para la unión de PALA.
7. Hemos combinado análisis estructurales y mutagénicos para identificar los elementos estructurales específicos responsables de la comunicación y cooperatividad entre las subunidades de huATCasa.
8. Hemos producido y caracterizado el mutante de huATCasa R2024Q, causante de la primera y única enfermedad humana descrita asociada a CAD. Proponemos que el residuo R2024 es parte de un mecanismo conformacional que controla los cambios producidos en las subunidades tras la unión de los ligandos.
9. Hemos producido la construcción bifuncional que contiene los dominios DHOasa y ATCasa de CAD humana y hemos demostrado que estos forman el núcleo de hexamerización de las partículas de CAD.

REFERENCE LIST

REFERENCE LIST

ACKERMANN, W. W. & POTTER, V. R. 1949. Enzyme inhibition in relation to chemotherapy. *Experimental Biology and Medicine,* 72**,** 1-9.

ADAMS, P. D., AFONINE, P. V., BUNKÓCZI, G., CHEN, V. B., DAVIS, I. W., ECHOLS, N., HEADD, J. J., HUNG, L. W., KAPRAL, G. J. & GROSSE-KUNSTLEVE, R. W. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography,* 66**,** 213-221.

AHUJA, A., PURCAREA, C., EBERT, R., SADECKI, S., GUY, H. I. & EVANS, D. R. 2004. Aquifex aeolicus dihydroorotase: association with aspartate transcarbamoylase switches on catalytic activity. *J Biol Chem,* 279**,** 53136-44.

ALLEWELL, N. M. 1989. Escherichia coli aspartate transcarbamoylase: structure, energetics, and catalytic and regulatory mechanisms. *Annu Rev Biophys Biophys Chem,* 18**,** 71-92.

ALLEWELL, N. M. & LICATA, V. J. 1995. [28] Thermodynamic approaches to understanding aspartate transcarbamylase. *Methods in enzymology,* 259**,** 608-628.

ANDERSON, P. M. & MEISTER, A. 1966. Control of Escherichia coli Carbamyl Phosphate Synthetase by Purine and Pyrimidine Nucleotides*. *Biochemistry,* 5**,** 3164-3169.

AOKI, T. & WEBER, G. 1981. Carbamoyl phosphate synthetase (glutamine-hydrolyzing): increased activity in cancer cells. *Science,* 212**,** 463-465.

ARICESCU, A. R., LU, W. & JONES, E. Y. 2006. A time-and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallographica Section D: Biological Crystallography,* 62**,** 1243-1250.

ASAI, T., O'SULLIVAN, W. J., KOBAYASHI, M., GERO, A. M., YOKOGAWA, M. & TATIBANA, M. 1983. Enzymes of the de novo pyrimidine biosynthetic pathway in Toxoplasma gondii. *Molecular and biochemical parasitology,* 7**,** 89-100.

BARBER, M. C., PRICE, N. T. & TRAVERS, M. T. 2005. Structure and regulation of acetyl-CoA carboxylase genes of metazoa. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids,* 1733**,** 1-28.

BEN-SAHRA, I., HOWELL, J. J., ASARA, J. M. & MANNING, B. D. 2013. Stimulation of de novo pyrimidine synthesis by growth signaling through mTOR and S6K1. *Science,* 339**,** 1323-8.

BERG, J. M., TYMOCZKO, J. L. & STRYER, L. 2002. In de Novo Synthesis, the Pyrimidine Ring Is Assembled from Bicarbonate, Aspartate, and Glutamine.

BERROW, N. S., ALDERTON, D., SAINSBURY, S., NETTLESHIP, J., ASSENBERG, R., RAHMAN, N., STUART, D. I. & OWENS, R. J. 2007. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res,* 35**,** e45.

BETHELL, M. R. & JONES, M. E. 1969. Molecular size and feedback-regulation characteristics of bacterial aspartate transcarbamylases. *Archives of biochemistry and biophysics,* 134**,** 352-365.

BOJANOWSKI, R., GAUDY, E., VALENTINE, R. C. & WOLFE, R. S. 1964. Oxamic transcarbamylase of Streptococcus allantoicus. *Journal of bacteriology,* 87**,** 75-80.

BOYD, K. E. & FARNHAM, P. J. 1997. Myc versus USF: discrimination at the cad gene is determined by core promoter elements. *Molecular and Cellular Biology,* 17**,** 2529-2537.

BRABSON, J. S. & SWITZER, R. L. 1975. Purification and properties of Bacillus subtilis aspartate transcarbamylase. *Journal of Biological Chemistry,* 250**,** 8664-8669.

BRADFORD, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry,* 72**,** 248-254.

BREMER, H. & DENNIS, P. P. 1996. Escherichia coli and Salmonella: cellular and molecular biology. *Washington (DC): American Society for Microbiology. Chapter, Modulation of chemical composition and other parameters of the cell by growth rate***,** 1553-1569.

BÜLTER, T. & ELLING, L. 1999. Enzymatic synthesis of nucleotide sugars. *Glycoconjugate journal,* 16**,** 147-159.

BUSH, A., MATEYAK, M., DUGAN, K., OBAYA, A., ADACHI, S., SEDIVY, J. & COLE, M. 1998. c-myc null cells misregulate cad and gadd45 but not other proposed c-Myc targets. *Genes & development,* 12**,** 3797-3802.

CALVA, E., LOWENSTEIN, J. M. & COHEN, P. P. 1959. Carbamyl phosphate-aspartate transcarbamylase activity in tumors. *Cancer research,* 19**,** 101-103.

CARREY, E. A. 1995. The shape of CAD. *Paths Pyrimidines,* 3**,** 68-72.

CARREY, E. A., CAMPBELL, D. G. & HARDIE, D. G. 1985. Phosphorylation and activation of hamster carbamyl phosphate synthetase II by cAMP-dependent protein kinase. A novel mechanism for regulation of pyrimidine nucleotide biosynthesis. *EMBO J,* 4**,** 3735-42.

CARREY, E. A. & HARDIE, D. G. 1988. Mapping of catalytic domains and phosphorylation sites in the multifunctional pyrimidine-biosynthetic protein CAD. *Eur J Biochem,* 171**,** 583-8.

CHANGEUX, J.-P. & RUBIN, M. M. 1968. Allosteric interactions in aspartate transcarbamylase. III. Interpretation of experimental data in terms of the model of Monod, Wyman, and Changeux. *Biochemistry,* 7**,** 553-560.

CHEN, P., VAN VLIET, F., VAN DE CASTEELE, M., LEGRAIN, C., CUNIN, R. & GLANSDORFF, N. 1998. Aspartate transcarbamylase from the hyperthermophilic eubacterium Thermotoga maritima: fused catalytic and regulatory polypeptides form an allosteric enzyme. *Journal of bacteriology,* 180**,** 6389-6391.

CHEN, V. B., ARENDALL, W. B., HEADD, J. J., KEEDY, D. A., IMMORMINO, R. M., KAPRAL, G. J., MURRAY, L. W., RICHARDSON, J. S. & RICHARDSON, D. C. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography,* 66**,** 12-21.

CHRISTOPHERSON, R. I. & DUGGLEBY, R. G. 1983. Metabolic resistance: the protection of enzymes against drugs which are tight‐binding inhibitors by the accumulation of substrate. *European Journal of Biochemistry,* 134**,** 331-335.

CHRISTOPHERSON, R. I. & JONES, M. E. 1980. The overall synthesis of L-5, 6-dihydroorotate by multienzymatic protein pyr1-3 from hamster cells. Kinetic studies, substrate channeling, and the effects of inhibitors. *Journal of Biological Chemistry,* 255**,** 11381-11395.

CHRISTOPHERSON, R. I. & LYONS, S. D. 1990. Potent inhibitors of de novo pyrimidine and purine biosynthesis as chemotherapeutic agents. *Medicinal research reviews,* 10**,** 505-548.

CHRISTOPHERSON, R. I., LYONS, S. D. & WILSON, P. K. 2002. Inhibitors of de novo nucleotide biosynthesis as drugs. *Accounts of chemical research,* 35**,** 961-971.

COHEN, R. E. & SCHACHMAN, H. K. 1986. Kinetics of the interaction of N-(phosphonacetyl)-L-aspartate with the catalytic subunit of aspartate transcarbamoylase. A slow conformational change subsequent to binding. *J Biol Chem,* 261**,** 2623-31.

COLEMAN, P. F., SUTTLE, D. P. & STARK, G. R. 1977. Purification from hamster cells of the multifunctional protein that initiates de novo synthesis of pyrimidine nucleotides. *J Biol Chem,* 252**,** 6379-85.

COLLINS, K. D. & STARK, G. R. 1971. Aspartate transcarbamylase interaction with the transition state analogue N-(phosphonacetyl)-L-aspartate. *Journal of Biological Chemistry,* 246**,** 6599-6605.

CONNER, T. W. & RAWLS JR, J. M. 1982. Analysis of the phenotypes exhibited by rudimentary-like mutants of Drosophila melanogaster. *Biochemical genetics,* 20**,** 607-619.

CONNOLLY, G. P. & DULEY, J. A. 1999. Uridine and its nucleotides: biological actions, therapeutic potentials. *Trends in pharmacological sciences,* 20**,** 218-225.

COOPER, G. M. 2000. Pathways of intracellular signal transduction.

COSTA, S., ALMEIDA, A., CASTRO, A. & DOMINGUES, L. 2014. Fusion tags for protein solubility, purification, and immunogenicity in Escherichia coli: the novel Fh8 system. *Recombinant protein expression in microbial systems***,** 24.

COX, J. A., LAMORA, A., JOHNSON, S. L. & VOIGT, M. M. 2014. Novel role for carbamoyl phosphate synthetase 2 in cranial sensory circuit formation. *International Journal of Developmental Neuroscience,* 33**,** 41-48.

DAVIDSON, J. N., CHEN, K. C., JAMISON, R. S., MUSMANNO, L. A. & KERN, C. B. 1993. The evolutionary history of the first three enzymes in pyrimidine biosynthesis. *Bioessays,* 15**,** 157-164.

DAVIDSON, J. N. & NISWANDER, L. A. 1983. Partial cDNA sequence to a hamster gene corrects defect in Escherichia coli pyrB mutant. *Proceedings of the National Academy of Sciences,* 80**,** 6897-6901.

DAVIDSON, J. N. & PATTERSON, D. 1979. Alteration in structure of multifunctional protein from Chinese hamster ovary cells defective in pyrimidine biosynthesis. *Proceedings of the National Academy of Sciences,* 76**,** 1731-1735.

DAVIDSON, J. N., RUMSBY, P. C. & TAMAREN, J. 1981. Organization of a multifunctional protein in pyrimidine biosynthesis. Analyses of active, tryptic fragments. *J Biol Chem,* 256**,** 5220-5.

DE CIMA, S., POLO, L. M., DÍEZ-FERNÁNDEZ, C., MARTÍNEZ, A. I., CERVERA, J., FITA, I. & RUBIO, V. 2015. Structure of human carbamoyl phosphate synthetase: deciphering the on/off switch of human ureagenesis. *Scientific reports,* 5.

DENIS-DUPHIL, M. & KAPLAN, J. G. 1976. Fine structure of the URA2 locus in Saccharomyces cerevisiae. *Molecular and General Genetics MGG,* 145**,** 259-271.

DOUETTE, P., NAVET, R., GERKENS, P., GALLENI, M., LÉVY, D. & SLUSE, F. E. 2005. Escherichia coli fusion carrier proteins act as solubilizing agents for recombinant uncoupling protein 1 through interactions with GroEL. *Biochemical and biophysical research communications,* 333**,** 686-693.

DUNN, M. F., NIKS, D., NGO, H., BARENDS, T. R. M. & SCHLICHTING, I. 2008. Tryptophan synthase: the workings of a channeling nanomachine. *Trends in biochemical sciences,* 33**,** 254-264.

DUTTA, S., WHICHER, J. R., HANSEN, D. A., HALE, W. A., CHEMLER, J. A., CONGDON, G. R., NARAYAN, A. R., HÅKANSSON, K., SHERMAN, D. H. & SMITH, J. L. 2014. Structure of a modular polyketide synthase. *Nature,* 510**,** 512.

ELSE, A. J. & HERVÉ, G. 1990. A microtiter plate assay for aspartate transcarbamylase. *Analytical biochemistry,* 186**,** 219-221.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography,* 66**,** 486-501.

ENGLAND, P., LECONTE, C., TAUC, P. & HERVE, G. 1994. Apparent cooperativity for carbamoylphosphate in Escherichia coli aspartate transcarbamoylase only reflects cooperativity for aspartate. *Eur J Biochem,* 222**,** 775-80.

EVANS, D. R. 1986. CAD, a chimeric protein that initiates de novo pyrimidine biosynthesis in higher eukaryotes. *Multidomain proteins—structure and evolution. Elsevier Biomedical Press, Amsterdam***,** 283-331.

EVANS, D. R. & GUY, H. I. 2004. Mammalian pyrimidine biosynthesis: fresh insights into an ancient pathway. *J Biol Chem,* 279**,** 33035-8.

FLORES, M. V. C., ATKINS, D., WADE, D., O'SULLIVAN, W. J. & STEWART, T. S. 1997. Inhibition of Plasmodium falciparum Proliferationin Vitro by Ribozymes. *Journal of Biological Chemistry,* 272**,** 16940-16945.

FOX, B. A. & BZIK, D. J. 2002. De novo pyrimidine biosynthesis is required for virulence of Toxoplasma gondii. *Nature,* 415**,** 926-929.

FOX, J. D., KAPUST, R. B. & WAUGH, D. S. 2001. Single amino acid substitutions on the surface of Escherichia coli maltose‐binding protein can have a profound impact on the solubility of fusion proteins. *Protein Science,* 10**,** 622-630.

FRANKS, D. M., IZUMIKAWA, T., KITAGAWA, H., SUGAHARA, K. & OKKEMA, P. G. 2006. C. elegans pharyngeal morphogenesis requires both de novo synthesis of pyrimidines and synthesis of heparan sulfate proteoglycans. *Developmental biology,* 296**,** 409-420.

GERHART, J. C. & HOLOUBEK, H. 1967. The purification of aspartate transcarbamylase of Escherichia coli and separation of its protein subunits. *Journal of Biological Chemistry,* 242**,** 2886-2892.

GERHART, J. C. & PARDEE, A. B. 1962. The enzymology of control by feedback inhibition. *J Biol Chem,* 237**,** 891-896.

GERHART, J. C. & PARDEE, A. B. 1963. The Effect of the Feedback Inhibitor, CTP, on Subunit Interactions in Aspartate Transcarbamylase. *Cold Spring Harbor Symposia on Quantitative Biology,* 28**,** 491-496.

GERHART, J. C. & SCHACHMAN, H. K. 1965. Distinct subunits for the regulation and catalytic activity of aspartate transcarbamylase. *Biochemistry,* 4**,** 1054-62.

GERO, A. M., BROWN, G. V. & O'SULLIVAN, W. J. 1984. Pyrimidine de novo synthesis during the life cycle of the intraerythrocytic stage of Plasmodium falciparum. *The Journal of parasitology***,** 536-541.

GIBBONS, I., YANG, Y. R. & SCHACHMAN, H. K. 1974. Cooperative interactions in aspartate transcarbamoylase. 1. Hybrids composed of native and chemically inactivated catalytic polypeptide chains. *Proc Natl Acad Sci U S A,* 71**,** 4452-6.

GOUAUX, J. E., KRAUSE, K. L. & LIPSCOMB, W. N. 1987. The catalytic mechanism of Escherichia coli aspartate carbamoyltransferase: a molecular modelling study. *Biochemical and biophysical research communications,* 142**,** 893-897.

GOUAUX, J. E., STEVENS, R. C. & LIPSCOMB, W. N. 1990. Crystal structures of aspartate carbamoyltransferase ligated with phosphonoacetamide, malonate, and CTP or ATP at 2.8-. ANG. resolution and neutral pH. *Biochemistry,* 29**,** 7702-7715.

GRANDE-GARCÍA, A., LALLOUS, N., DÍAZ-TEJADA, C. & RAMÓN-MAIQUES, S. 2014. Structure, functional characterization, and evolution of the dihydroorotase domain of human CAD. *Structure,* 22**,** 185-198.

GRAVES, L. M., GUY, H. I., KOZLOWSKI, P., HUANG, M., LAZAROWSKI, E., POPE, R. M., COLLINS, M. A., DAHLSTRAND, E. N., EARP, H. S., 3RD & EVANS, D. R. 2000. Regulation of carbamoyl phosphate synthetase by MAP kinase. *Nature,* 403**,** 328-32.

GRAYSON, D. R. & EVANS, D. R. 1983. The isolation and characterization of the aspartate transcarbamylase domain of the multifunctional protein, CAD. *Journal of Biological Chemistry,* 258**,** 4123-4129.

GREM, J. L., KING, S. A., O'DWYER, P. J. & LEYLAND-JONES, B. 1988. Biochemistry and clinical activity of N-(phosphonacetyl)-L-aspartate: a review. *Cancer Res,* 48**,** 4441-54.

GUYONVARCH, A., NGUYEN-JUILLERET, M., HUBERT, J. C. & LACROUTE, F. 1988. Structure of the Saccharomyces cerevisiae URA4 gene encoding dihydroorotase. *Molecular and General Genetics MGG,* 212**,** 134-141.

HANSEN, H. A. & EMBORG, C. 1994. Extra-and intracellular amino acid concentrations in continuous Chinese hamster ovary cell culture. *Applied microbiology and biotechnology,* 41**,** 560-564.

HAWKINS, A. R., LAMB, H. K., MOORE, J. D., CHARLES, I. G. & ROBERTS, C. F. 1993. Review Article: The pre-chorismate (shikimate) and quinate pathways in filamentous fungi: theoretical and practical aspects. *Microbiology,* 139**,** 2891-2899.

HEMMENS, B. & CARREY, E. A. 1994. Proteolytic cleavage of the multienzyme polypeptide CAD to release the mammalian aspartate transcarbamoylase. Biochemical comparison with the homologous Escherichia coli catalytic subunit. *Eur J Biochem,* 225**,** 845-53.

HO, S. N., HUNT, H. D., HORTON, R. M., PULLEN, J. K. & PEASE, L. R. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene,* 77**,** 51-59.

HONZATKO, R. B., CRAWFORD, J. L., MONACO, H. L., LADNER, J. E., EWARDS, B. F. P., EVANS, D. R., WARREN, S. G., WILEY, D. C., LADNER, R. C. & LIPSCOMB, W. N. 1982. Crystal and molecular structures of native and CTP-liganded aspartate carbamoyltransferase from Escherichia coli. *Journal of molecular biology,* 160**,** 219-263.

HOOGENRAAD, N. J. & LEE, D. C. 1974. Effect of uridine on de novo pyrimidine biosynthesis in rat hepatoma cells in culture. *Journal of Biological Chemistry,* 249**,** 2763-2768.

HOOGENRAAD, N. J., LEVINE, R. L. & KRETCHMER, N. 1971. Copurification of carbamoyl phosphate synthetase and aspartate transcarbamoylase from mouse spleen. *Biochem Biophys Res Commun,* 44**,** 981-8.

HOWLETT, G. J. & SCHACHMAN, H. K. 1977. Allosteric regulation of aspartate transcarbamoylase. Changes in the sedimentation coefficient promoted by the bisubstrate analogue N-(phosphonacetyl)-L-aspartate. *Biochemistry,* 16**,** 5077-83.

HUANG, D. T. C., THOMAS, M. A. W. & CHRISTOPHERSON, R. I. 1999. Divalent metal derivatives of the hamster dihydroorotase domain. *Biochemistry,* 38**,** 9964-9970.

HUANG, M. & GRAVES, L. M. 2003. De novo synthesis of pyrimidine nucleotides; emerging interfaces with signal transduction pathways. *Cellular and Molecular Life Sciences CMLS,* 60**,** 321-336.

INOUE, H., NOJIMA, H. & OKAYAMA, H. 1990. High efficiency transformation of Escherichia coli with plasmids. *Gene,* 96**,** 23-28.

IRVINE, H. S., SHAW, S. M., PATON, A. & CARREY, E. A. 1997. A reciprocal allosteric mechanism for efficient transfer of labile intermediates between active sites in CAD, the mammalian pyrimidine-biosynthetic multienzyme polypeptide. *Eur J Biochem,* 247**,** 1063-73.

JIN, L., STEC, B., LIPSCOMB, W. N. & KANTROWITZ, E. R. 1999. Insights into the mechanisms of catalysis and heterotropic regulation of Escherichia coli aspartate transcarbamoylase based upon a structure of the enzyme complexed with the bisubstrate analogue N-phosphonacetyl-L-aspartate at 2.1 A. *Proteins,* 37**,** 729-42.

JOHNSON, R. K., INOUYE, T., GOLDIN, A. & STARK, G. R. 1976. Antitumor activity of N-(phosphonacetyl)-L-aspartic acid, a transition-state inhibitor of aspartate transcarbamylase. *Cancer research,* 36**,** 2720-2725.

JONES, M. E. 1980. Pyrimidine nucleotide biosynthesis in animals: genes, enzymes, and regulation of UMP biosynthesis. *Annual review of biochemistry,* 49**,** 253-279.

KABSCH, W. 2010. Xds. *Acta Crystallographica Section D: Biological Crystallography,* 66**,** 125-132.

KANTROWITZ, E. R. 2012. Allostery and cooperativity in Escherichia coli aspartate transcarbamoylase. *Archives of biochemistry and biophysics,* 519**,** 81-90.

KANTROWITZ, E. R. & LIPSCOMB, W. N. 1988. Escherichia coli aspartate transcarbamylase: the relation between structure and function. *Science,* 241**,** 669-674.

KANTROWITZ, E. R. & LIPSCOMB, W. N. 1990. Escherichia coli aspartate transcarbamoylase: the molecular basis for a concerted allosteric transition. *Trends Biochem Sci,* 15**,** 53-9.

KAPUST, R. B. & WAUGH, D. S. 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science,* 8**,** 1668-1674.

KELLY, R. E., MALLY, M. I. & EVANS, D. R. 1986. The dihydroorotase domain of the multifunctional protein CAD. Subunit structure, zinc content, and kinetics. *Journal of Biological Chemistry,* 261**,** 6073-6083.

KEMPE, T. D., SWYRYD, E. A., BRUIST, M. & STARK, G. R. 1976. Stable mutants of mammalian cells that overproduce the first three enzymes of pyrimidine nucleotide biosynthesis. *Cell,* 9**,** 541-550.

KENSLER, T. W., ERLICHMAN, C., JAYARAM, H. N., TYAGI, A. K., ARDALAN, B. & COONEY, D. A. 1980. Peripheral leukocytes as indicators of the enzymatic effects of N-(phosphonacetyl)-L-aspartic acid (PALA) on human L-aspartate transcarbamoylase (ATCase) activity. *Cancer Treat Rep,* 64**,** 967-73.

KIM, H., KELLY, R. E. & EVANS, D. R. 1992. The structural organization of the hamster multifunctional protein CAD. Controlled proteolysis, domains, and linkers. *J Biol Chem,* 267**,** 7177-84.

LABEDAN, B., BOYEN, A., BAETENS, M., CHARLIER, D., CHEN, P., CUNIN, R., DURBECO, V., GLANSDORFF, N., HERVE, G. & LEGRAIN, C. 1999. The evolutionary history of carbamoyltransferases: a complex set of paralogous genes was already present in the last universal common ancestor. *Journal of molecular evolution,* 49**,** 461-473.

LABEDAN, B., XU, Y., NAUMOFF, D. G. & GLANSDORFF, N. 2004. Using quaternary structures to assess the evolutionary history of proteins: the case of the aspartate carbamoyltransferase. *Molecular biology and evolution,* 21**,** 364-373.

LALLOUS, N., GRANDE-GARCIA, A., MOLINA, R. & RAMON-MAIQUES, S. 2012. Expression, purification, crystallization and preliminary X-ray diffraction analysis of the dihydroorotase domain of human CAD. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 68**,** 1341-5.

LANE, A. N. & FAN, T. W. M. 2015. Regulation of mammalian nucleotide metabolism and biosynthesis. *Nucleic acids research***,** gkv047.

LEE, L., KELLY, R. E., PASTRA-LANDIS, S. C. & EVANS, D. R. 1985. Oligomeric structure of the multifunctional protein CAD that initiates pyrimidine biosynthesis in mammalian cells. *Proc Natl Acad Sci U S A,* 82**,** 6802-6.

LEGRAIN, C., DEMAREZ, M., GLANSDORFF, N. & PIÉRARD, A. 1995. Ammonia-dependent synthesis and metabolic channelling of carbamoyl phosphate in the hyperthermophilic archaeon Pyrococcus furiosus. *Microbiology,* 141**,** 1093-1099.

LICATA, V. J. & ALLEWELL, N. M. 1997. Is substrate inhibition a consequence of allostery in aspartate transcarbamylase? *Biophys Chem,* 64**,** 225-34.

LIPSCOMB, W. N. & KANTROWITZ, E. R. 2012. Structure and mechanisms of Escherichia coli aspartate transcarbamoylase. *Acc Chem Res,* 45**,** 444-53.

LIU, X., GUY, H. I. & EVANS, D. R. 1994. Identification of the regulatory domain of the mammalian multifunctional protein CAD by the construction of an Escherichia coli hamster hybrid carbamyl-phosphate synthetase. *J Biol Chem,* 269**,** 27747-55.

LODISH, H., BERK, A., ZIPURSKY, S. L., MATSUDAIRA, P., BALTIMORE, D. & DARNELL, J. 2000. Processing of rRNA and tRNA. *Molecular Cell Biology. 4th Edition*.

LÖFFLER, M., FAIRBANKS, L. D., ZAMEITAT, E., MARINAKI, A. M. & SIMMONDS, H. A. 2005. Pyrimidine pathways in health and disease. *Trends in molecular medicine,* 11**,** 430-437.

LUE, P. F. & KAPLAN, J. G. 1969. The aspartate transcarbamylase and carbamoyl phosphate synthetase of yeast: a multi-functional enzyme complex. *Biochemical and biophysical research communications,* 34**,** 426-433.

MAJOR, J. G., JR., WALES, M. E., HOUGHTON, J. E., MALEY, J. A., DAVIDSON, J. N. & WILD, J. R. 1989. Molecular evolution of enzyme structure: construction of a hybrid hamster/Escherichia coli aspartate transcarbamoylase. *J Mol Evol,* 28**,** 442-50.

MALEY, J. A. & DAVIDSON, J. N. 1988. The aspartate transcarbamylase domain of a mammalian multifunctional protein expressed as an independent enzyme in Escherichia coli. *Mol Gen Genet,* 213**,** 278-84.

MALLY, M. I., GRAYSON, D. R. & EVANS, D. R. 1981. Controlled proteolysis of the multifunctional protein that initiates pyrimidine biosynthesis in mammalian cells: evidence for discrete structural domains. *Proceedings of the National Academy of Sciences,* 78**,** 6647-6651.

MARSH, J. A., HERNÁNDEZ, H., HALL, Z., AHNERT, S. E., PERICA, T., ROBINSON, C. V. & TEICHMANN, S. A. 2013. Protein complexes are under evolutionary selection to assemble via ordered pathways. *Cell,* 153**,** 461-470.

MCCOY, A. J., GROSSE-KUNSTLEVE, R. W., ADAMS, P. D., WINN, M. D., STORONI, L. C. & READ, R. J. 2007. Phaser crystallographic software. *Journal of applied crystallography,* 40**,** 658-674.

MENDES, K. R. & KANTROWITZ, E. R. 2010. A cooperative Escherichia coli aspartate transcarbamoylase without regulatory subunits. *Biochemistry,* 49**,** 7694-703.

MONOD, J., WYMAN, J. & CHANGEUX, J. P. 1965. ON THE NATURE OF ALLOSTERIC TRANSITIONS: A PLAUSIBLE MODEL. *J Mol Biol,* 12**,** 88-118.

MORENO-BELTRÁN, B., DÍAZ-MORENO, I., GONZÁLEZ-ARZOLA, K., GUERRA-CASTELLANO, A., VELÁZQUEZ-CAMPOY, A., MIGUEL, A. & DÍAZ-QUINTANA, A. 2015. Respiratory complexes III and IV can each bind two molecules of cytochrome c at low ionic strength. *FEBS letters,* 589**,** 476-483.

MORRISON, J. F. 1969. Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. *Biochim Biophys Acta,* 185**,** 269-86.

MORRISON, J. F. 1982. The slow-binding and slow, tight-binding inhibition of enzyme-catalysed reactions. *Trends in Biochemical Sciences,* 7**,** 102-105.

MURSHUDOV, G. N., SKUBÁK, P., LEBEDEV, A. A., PANNU, N. S., STEINER, R. A., NICHOLLS, R. A., WINN, M. D., LONG, F. & VAGIN, A. A. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography,* 67**,** 355-367.

NEWELL, J. O., MARKBY, D. W. & SCHACHMAN, H. K. 1989. Cooperative binding of the bisubstrate analog N-(phosphonacetyl)-L-aspartate to aspartate transcarbamoylase and the heterotropic effects of ATP and CTP. *J Biol Chem,* 264**,** 2476-81.

NG, B. G., WOLFE, L. A., ICHIKAWA, M., MARKELLO, T., HE, M., TIFFT, C. J., GAHL, W. A. & FREEZE, H. H. 2015. Biallelic mutations in CAD, impair de novo pyrimidine biosynthesis and decrease glycosylation precursors. *Hum Mol Genet,* 24**,** 3050-7.

NG, S. B., BUCKINGHAM, K. J., LEE, C., BIGHAM, A. W., TABOR, H. K., DENT, K. M., HUFF, C. D., SHANNON, P. T., JABS, E. W. & NICKERSON, D. A. 2010. Exome sequencing identifies the cause of a mendelian disorder. *Nature genetics,* 42**,** 30-35.

NIESEN, F. H., BERGLUND, H. & VEDADI, M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc,* 2**,** 2212-21.

NØRBY, S. 1970. A specific nutritional requirement for pyrimidines in rudimentary mutants of Drosophila melanogaster. *Hereditas,* 66**,** 205-214.

NUZUM, C. T. & SNODGRASS, P. J. 1976. Multiple assays of the five urea-cycle enzymes in human liver homogenates. *The urea cycle***,** 325-349.

NYUNOYA, H., BROGLIE, K. E., WIDGREN, E. E. & LUSTY, C. J. 1985. Characterization and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase I of rat. *Journal of Biological Chemistry,* 260**,** 9346-9356.

O'DWYER, P. 1990. The role of low-dose PALA in biochemical modulation. *Pharmacology & therapeutics,* 48**,** 371-380.

O'DWYER, P. J., PAUL, A. R., WALCZAK, J., WEINER, L. M., LITWIN, S. & COMIS, R. L. 1990. Phase II study of biochemical modulation of fluorouracil by low-dose PALA in patients with colorectal cancer. *Journal of Clinical Oncology,* 8**,** 1497-1503.

PADILLA, J. E. & YEATES, T. O. 2003. A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallographica Section D: Biological Crystallography,* 59**,** 1124-1130.

PASTRA-LANDIS, S. C., EVANS, D. R. & LIPSCOMB, W. N. 1978. The effect of pH on the cooperative behavior of aspartate transcarbamylase from Escherichia coli. *J Biol Chem,* 253**,** 4624-30.

PERUTZ, M. F. 1989. Mechanisms of cooperativity and allosteric regulation in proteins. *Quarterly reviews of biophysics,* 22**,** 139-237.

POLO, L. M., GIL-ORTIZ, F., CANTÍN, A. & RUBIO, V. 2012. New insight into the transcarbamylase family: the structure of putrescine transcarbamylase, a key catalyst for fermentative utilization of agmatine. *PloS one,* 7**,** e31528.

PORTER, R. W., MODEBE, M. O. & STARK, G. R. 1969. Aspartate transcarbamylase. Kinetic studies of the catalytic subunit. *J Biol Chem,* 244**,** 1846-59.

PRESCOTT, L. M. & JONES, M. E. 1969. Modified methods for the determination of carbamyl aspartate. *Analytical biochemistry,* 32**,** 408-419.

QIU, Y. & DAVIDSON, J. N. 1998. Aspartate-90 and arginine-269 of hamster aspartate transcarbamylase affect the oligomeric state of a chimaeric protein with an Escherichia coli maltose-binding domain. *Biochem J,* 329 (Pt 2)**,** 243-7.

QIU, Y. & DAVIDSON, J. N. 2000. Substitutions in the aspartate transcarbamoylase domain of hamster CAD disrupt oligomeric structure. *Proceedings of the National Academy of Sciences,* 97**,** 97-102.

RADZICKA, A. & WOLFENDEN, R. 1995. A proficient enzyme. *Science,* 267**,** 90.

RAWLS, J., KNECHT, W., DIEKERT, K., LILL, R. & LÖFFLER, M. 2000. Requirements for the mitochondrial import and localization of dihydroorotate dehydrogenase. *European Journal of biochemistry,* 267**,** 2079-2087.

ROBITAILLE, A. M., CHRISTEN, S., SHIMOBAYASHI, M., CORNU, M., FAVA, L. L., MOES, S., PRESCIANOTTO-BASCHONG, C., SAUER, U., JENOE, P. & HALL, M. N. 2013. Quantitative phosphoproteomics reveal mTORC1 activates de novo pyrimidine synthesis. *Science,* 339**,** 1320-3.

ROSVOLD, E., SCHILDER, R., WALCZAK, J., DIFINO, S. M., FLYNN, P. J., BANERJEE, T. K., HEIM, W. J., ENGSTROM, P. E., OZOLS, R. F. & O'DWYER, P. J. 1992. Phase II trial of PALA in combination with 5-fluorouracil in advanced pancreatic cancer. *Cancer chemotherapy and pharmacology,* 29**,** 305-308.

RUIZ‐RAMOS, A., GRANDE‐GARCÍA, A. & RAMÓN‐MAIQUES, S. Dihydroorotase Domain of Human CAD. *Encyclopedia of Inorganic and Bioinorganic Chemistry*.

SAHAY, N., GUY, H. I., LIU, X. & EVANS, D. R. 1998. Regulation of an Escherichia coli/mammalian chimeric carbamoyl-phosphate synthetase. *Journal of Biological Chemistry,* 273**,** 31195-31202.

SCHACHMAN, H. K. 1987. From allostery to mutagenesis: 20 years with aspartate transcarbamoylase. *Biochem Soc Trans,* 15**,** 772-5.

SCHUCK, P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal,* 78**,** 1606-1619.

SCHWENGER, B., SCHÖBER, S. & SIMON, D. 1993. DUMPS cattle carry a point mutation in the uridine monophosphate synthase gene. *Genomics,* 16**,** 241-244.

SCULLY, J. L. & EVANS, D. R. 1991. Comparative modeling of mammalian aspartate transcarbamylase. *Proteins,* 9**,** 191-206.

SEYMOUR, K. K., LYONS, S. D., PHILLIPS, L., RIECKMANN, K. H. & CHRISTOPHERSON, R. I. 1994. Cytotoxic effects of inhibitors of de novo pyrimidine biosynthesis upon Plasmodium falciparum. *Biochemistry,* 33**,** 5268-5274.

SHI, D., ALLEWELL, N. M. & TUCHMAN, M. 2015. From Genome to Structure and Back Again: A Family Portrait of the Transcarbamylases. *International journal of molecular sciences,* 16**,** 18836-18864.

SHOAF, W. T. & JONES, M. E. 1971. Initial steps in pyrimidine synthesis in Ehrlich ascites carcinoma. *Biochemical and biophysical research communications,* 45**,** 796-802.

SIGOILLOT, F. D., BERKOWSKI, J. A., SIGOILLOT, S. M., KOTSIS, D. H. & GUY, H. I. 2003. Cell cycle-dependent regulation of pyrimidine biosynthesis. *J Biol Chem,* 278**,** 3403-9.

SIGOILLOT, F. D., EVANS, D. R. & GUY, H. I. 2002. Autophosphorylation of the mammalian multifunctional protein that initiates de novo pyrimidine biosynthesis. *J Biol Chem,* 277**,** 24809-17.

SIGOILLOT, F. D., SIGOILLOT, S. M. & GUY, H. I. 2004. Breakdown of the regulatory control of pyrimidine biosynthesis in human breast cancer cells. *Int J Cancer,* 109**,** 491-8.

SIMMER, J. P., KELLY, R. E., RINKER, A. G., JR., SCULLY, J. L. & EVANS, D. R. 1990. Mammalian carbamyl phosphate synthetase (CPS). DNA sequence and evolution of the CPS domain of the Syrian hamster multifunctional protein CAD. *J Biol Chem,* 265**,** 10395-402.

SIMMER, J. P., KELLY, R. E., SCULLY, J. L., GRAYSON, D. R., RINKER, A. G., JR., BERGH, S. T. & EVANS, D. R. 1989. Mammalian aspartate transcarbamylase (ATCase): sequence of the ATCase domain and interdomain linker in the CAD multifunctional polypeptide and properties of the isolated domain. *Proc Natl Acad Sci U S A,* 86**,** 4382-6.

SMITH, S. 1994. The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. *The FASEB journal,* 8**,** 1248-1259.

SOUCIET, J. L., NAGY, M., LE GOUAR, M., LACROUTE, F. & POTIER, S. 1989. Organization of the yeast URA2 gene: identification of a defective dihydroorotase-like domain in the multifunctional carbamoylphosphate synthetase-aspartate transcarbamylase complex. *Gene,* 79**,** 59-70.

STEBBINS, J. W. & KANTROWITZ, E. R. 1989. The importance of the link between Glu204 of the catalytic chain and Arg130 of the regulatory chain for the homotropic and heterotropic properties of Escherichia coli aspartate transcarbamoylase. *J Biol Chem,* 264**,** 14860-4.

STEBBINS, J. W. & KANTROWITZ, E. R. 1992. Conversion of the noncooperative Bacillus subtilis aspartate transcarbamoylase into a cooperative enzyme by a single amino acid substitution. *Biochemistry,* 31**,** 2328-32.

STEBBINS, J. W., XU, W. & KANTROWITZ, E. R. 1989. Three residues involved in binding and catalysis in the carbamyl phosphate binding site of Escherichia coli aspartate transcarbamylase. *Biochemistry,* 28**,** 2592-2600.

STEVENS, R. C., REINISCH, K. M. & LIPSCOMB, W. N. 1991. Molecular structure of Bacillus subtilis aspartate transcarbamoylase at 3.0 A resolution. *Proc Natl Acad Sci U S A,* 88**,** 6087-91.

STUDIER, F. W. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein expression and purification,* 41**,** 207-234.

SUCHI, M., MIZUNO, H., KAWAI, Y., TSUBOI, T., SUMI, S., OKAJIMA, K., HODGSON, M. E., OGAWA, H. & WADA, Y. 1997. Molecular cloning of the human UMP synthase gene and characterization of point mutations in two hereditary orotic aciduria families. *American journal of human genetics,* 60**,** 525.

SWYRYD, E. A., SEAVER, S. S. & STARK, G. R. 1974. N-(phosphonacetyl)-L-aspartate, a potent transition state analog inhibitor of aspartate transcarbamylase, blocks proliferation of mammalian cells in culture. *J Biol Chem,* 249**,** 6945-50.

TRAMELL, P. R. & CAMPBELL, J. W. 1970. Carbamyl phosphate synthesis in a land snail, Strophocheilus oblongus. *Journal of Biological Chemistry,* 245**,** 6634-6641.

TRAN, T. H., HSIAO, Y.-S., JO, J., CHOU, C.-Y., DIETRICH, L. E. P., WALZ, T. & TONG, L. 2015. Structure and function of a single-chain, multi-domain long-chain acyl-CoA carboxylase. *Nature,* 518**,** 120-124.

TRAUT, T. W. & JONES, M. E. 1996. Uracil metabolism—UMP synthesis from orotic acid or uridine and conversion of uracil to β-alanine: enzymes and cDNAs. *Progress in nucleic acid research and molecular biology,* 53**,** 1-78.

TSUBOI, K. K., EDMUNDS, H. N. & KWONG, L. K. 1977. Selective inhibition of pyrimidine biosynthesis and effect on proliferative growth of colonic cancer cells. *Cancer research,* 37**,** 3080-3087.

WAHL, G. M., PADGETT, R. A. & STARK, G. R. 1979. Gene amplification causes overproduction of the first three enzymes of UMP synthesis in N-(phosphonacetyl)-L-aspartate-resistant hamster cells. *Journal of Biological Chemistry,* 254**,** 8679-8689.

WAKIL, S. J. 1989. Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry,* 28**,** 4523-4530.

WANG, J., STIEGLITZ, K. A., CARDIA, J. P. & KANTROWITZ, E. R. 2005. Structural basis for ordered substrate binding and cooperativity in aspartate transcarbamoylase. *Proc Natl Acad Sci U S A,* 102**,** 8881-6.

WEST, J. M., XIA, J., TSURUTA, H., GUO, W., O'DAY, E. M. & KANTROWITZ, E. R. 2008. Time evolution of the quaternary structure of Escherichia coli aspartate transcarbamoylase upon reaction with the natural substrates and a slow, tight-binding inhibitor. *J Mol Biol,* 384**,** 206-18.

WILLER, G. B., LEE, V. M., GREGG, R. G. & LINK, B. A. 2005. Analysis of the Zebrafish perplexed mutation reveals tissue-specific roles for de novo pyrimidine synthesis during development. *Genetics,* 170**,** 1827-1837.

WILLIAMS, L. G., BERNHARDT, S. & DAVIS, R. H. 1970. Copurification of pyrimidine-specific carbamyl phosphate synthetase and aspartate transcarbamylase of Neurospora crassa. *Biochemistry,* 9**,** 4329-4335.

WITKOWSKI, A., RANGAN, V. S., RANDHAWA, Z. I., AMY, C. M. & SMITH, S. 1991. Structural organization of the multifunctional animal fatty‐acid synthase. *European Journal of Biochemistry,* 198**,** 571-579.

WYATT, P. J. 1993. Light scattering and the absolute characterization of macromolecules. *Analytica chimica acta,* 272**,** 1-40.

YANG, S.-Y. & SCHULZ, H. 1983. The large subunit of the fatty acid oxidation complex from Escherichia coli is a multifunctional polypeptide. Evidence for the existence of a fatty acid oxidation operon (fad AB) in Escherichia coli. *Journal of Biological Chemistry,* 258**,** 9780-9785.

YEATES, T. O. 1997. [22] Detecting and overcoming crystal twinning. *Methods in enzymology,* 276**,** 344-358.

YOSHIDA, T., STARK, G. R. & HOOGENRAAD, N. J. 1974. Inhibition by N-(phosphonacetyl)-L-aspartate of aspartate transcarbamylase activity and drug-induced cell proliferation in mice. *Journal of Biological Chemistry,* 249**,** 6951-6955.

YUAN, X., LICATA, V. J. & ALLEWELL, N. M. 1996. Effects of assembly and mutations outside the active site on the functional pH dependence of Escherichia coli aspartate transcarbamylase. *Journal of Biological Chemistry,* 271**,** 1285-1294.

APPENDIXES